

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 17:23:01 ; Search time 187 Seconds
(without alignments)
126.879 Million cell updates/sec

Title: US-10-644-927-1
Perfect score: 304
Sequence: 1 KTYGTNGVHCTKNSLWGV.....GRLDILLGWATGAFGKTFH 54

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 304 | 100.0 | 54 | 9 AEB18127 | Aeb18127 Lactobaci |
| 2 | 69.5 | 22.9 | 420 | 6 ADA33974 | Ada33974 Acinetoba |
| 3 | 64 | 21.1 | 61 | 2 AAW17990 | Aaw17990 Sakacin P |
| 4 | 62.5 | 20.6 | 558 | 8 ADR89438 | Adr89438 crybun2-o |
| 5 | 62 | 20.4 | 42 | 2 AAW11633 | Aaw11633 Lactobaci |
| 6 | 60 | 19.7 | 37 | 2 AAW66446 | Aaw66446 Cationic |
| 7 | 60 | 19.7 | 37 | 3 AAY91745 | Aay91745 Cationic |
| 8 | 60 | 19.7 | 300 | 2 AAR56481 | Aar56481 CD38. 3/2 |
| 9 | 60 | 19.7 | 300 | 2 AAW37928 | Aaw37928 Amino aci |
| 10 | 60 | 19.7 | 300 | 4 AAB69069 | Aab69069 Human CD3 |
| 11 | 60 | 19.7 | 300 | 5 AAG61817 | Aag61817 Prostate |
| 12 | 60 | 19.7 | 300 | 6 ABU09711 | Abu09711 CD38 prot |
| 13 | 60 | 19.7 | 300 | 6 ABU09714 | Abu09714 CD38 prot |
| 14 | 60 | 19.7 | 300 | 7 ADD18904 | Add18904 Human dis |
| 15 | 60 | 19.7 | 300 | 7 ADD45262 | Add45262 Human Pro |
| 16 | 60 | 19.7 | 300 | 7 ADN40030 | Adn40030 Cancer/an |
| 17 | 60 | 19.7 | 300 | 8 ADM32177 | Adm32177 Human CD3 |
| 18 | 60 | 19.7 | 300 | 8 ADQ59526 | Adq59526 Human can |
| 19 | 60 | 19.7 | 300 | 8 ADQ88192 | Adq88192 Human 903 |
| 20 | 60 | 19.7 | 300 | 9 ADY15608 | Ady15608 PRO polyp |
| 21 | 60 | 19.7 | 300 | 9 ADY19608 | Ady19608 PRO polyp |
| 22 | 60 | 19.7 | 300 | 9 AD213907 | Ad213907 Human can |
| 23 | 60 | 19.7 | 688 | 4 AEG06883 | Aeg06883 Novel hum |
| 24 | 59.5 | 19.6 | 513 | 2 AAW79755 | Aaw79755 Euphorbia |

| | | | | | |
|----|------|------|-------|------------|--------------------|
| 25 | 59 | 19.4 | 44 | 2 AAR14564 | Aar14564 Bacterioc |
| 26 | 59 | 19.4 | 44 | 2 AAR91745 | Aar91745 Brevicin, |
| 27 | 59 | 19.4 | 62 | 2 AAR25078 | Aar25078 PA-1 bact |
| 28 | 59 | 19.4 | 300 | 6 ABU09712 | Abu09712 CD38 prot |
| 29 | 59 | 19.4 | 399 | 8 ADN23488 | Adn23488 Bacterial |
| 30 | 59 | 19.4 | 753 | 4 AAB50093 | Ab50093 GB1 prote |
| 31 | 57.5 | 18.9 | 461 | 7 ADC64559 | Adc64559 Trichodes |
| 32 | 57.5 | 18.9 | 461 | 8 ADQ07360 | Adq07360 Trichodes |
| 33 | 57.5 | 18.9 | 461 | 9 AEA00462 | Aea00462 Inorganic |
| 34 | 57.5 | 18.9 | 643 | 4 ABB71088 | Abb71088 Drosophil |
| 35 | 57 | 18.8 | 581 | 6 ABU44608 | Abu44608 Protein e |
| 36 | 57 | 18.8 | 1565 | 6 ABM69065 | Abm69065 Photorhab |
| 37 | 56.5 | 18.6 | 45 | 5 ABB05473 | Abb05473 Enterococ |
| 38 | 56.5 | 18.6 | 242 | 7 ADF04643 | Adf04643 Bacterial |
| 39 | 56.5 | 18.6 | 527 | 9 AEB28045 | Aeb28045 Bacteriop |
| 40 | 56 | 18.4 | 66 | 7 ADC95053 | Adc95053 E. faeciu |
| 41 | 56 | 18.4 | 6641 | 6 ABU42656 | Abu42656 Protein e |
| 42 | 56 | 18.4 | 10182 | 5 ABP38314 | Abp38314 Staphyloc |
| 43 | 56 | 18.4 | 10203 | 8 ADO84851 | Ado84851 S epiderm |
| 44 | 56 | 18.4 | 10203 | 8 ADO84803 | Ado84803 Staphyloc |
| 45 | 55.5 | 18.3 | 244 | 2 AAY06359 | Aay06359 Fusarium |

ALIGNMENTS

RESULT 1
AEB18127
ID AEB18127 standard; protein; 54 AA.
XX
AC AEB18127;
XX
DT 22-SEP-2005 (first entry)
XX
DE Lactobacillus salivarius strain PVD32 bacteriocin, OR7, SEQ ID NO:1.
XX
KW Bacteriocin; feedstuff; therapeutic; gene therapy; lactocin.
XX
OS Lactobacillus salivarius; strain PVD32.
XX
FN US2005153881-A1.
XX
PD 14-JUL-2005.
XX
PF 21-AUG-2003; 2003US-00644927.
XX
PR 21-AUG-2003; 2003US-00644927.
XX
PA (STER/) STERN N J.
PA (SVET/) SVETOCH E A.
PA (ERUS/) ERUSLANOV B V.
PA (VOLO/) VOLODINA L I.
PA (KOVA/) KOVALEV Y N.
PA (KUDR/) KUDRYAVTSEVA T Y.
PA (PERE/) PERELYGIN V V.
PA (POKH/) POKHILENKO V D.
PA (LEVCH/) LEVCHUK V P.
PA (BORZ/) BORZENKOV V N.
PA (SVET/) SVETOCH O E.
PA (MITS/) MITSEVICH E V.
PA (MITS/) MITSEVICH I P.
XX
PI Stern NJ, Svetoch EA, Eruslanov BV, Volodina LI, Kovalev YN;
PI Kudryavtseva TY, Perelygin VV, Pokhilenko VD, Levchuk VP;
PI Borzenkov VN, Svetoch OE, Mitseovich EV, Mitseovich IP;
XX
DR WPI; 2005-496854/50.
XX
PT New bacteriocin produced by a lactic acid producing bacterial strain
PT having the characteristics of strain NRRL B-30514, B-30510, B-30511 or B-
PT 30645, useful in producing therapeutic feed against bacterial
PT colonization in poultry.
XX

RESULT 4
ADR89438
ID ADR89438 standard; protein; 558 AA.

XX AC ADR89438;
XX DT 18-NOV-2004 (first entry)
XX DE crybun2-orf2.
XX KW delta-endotoxin; delta-endotoxin associate polypeptide;
XX KW expression cassette; transformation; transgenic; plant; bacteria;
XX KW lepidoptera; coleoptera; pest; pesticide; resistance;
XX KW pesticidal activity.

XX OS Bacillus thuringiensis.
XX PN W02004074462-A2.

XX PD 02-SEP-2004.
XX PF 20-FEB-2004; 2004WO-US005829.

XX PR 20-FEB-2003; 2003US-0448632P.
XX PR 20-FEB-2003; 2003US-0448633P.
XX PR 20-FEB-2003; 2003US-0448797P.
XX PR 20-FEB-2003; 2003US-0448806P.
XX PR 20-FEB-2003; 2003US-0448810P.
XX PR 20-FEB-2003; 2003US-0448812P.
XX PR 19-FEB-2004; 2004US-00781979.
XX PR 19-FEB-2004; 2004US-00782020.
XX PR 19-FEB-2004; 2004US-00782096.
XX PR 19-FEB-2004; 2004US-00782141.
XX PR 19-FEB-2004; 2004US-00782570.
XX PR 19-FEB-2004; 2004US-00783417.

XX PA (ATHE-) ATHENIX CORP.

XX PI Carozzi N, Hargliss T, Kozziel MG, Duck NB, Carr B;
XX WPI; 2004-635574/61.

XX PT New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids
XX PT and polypeptides, useful for killing lepidopteran or coleopteran pests or
XX PT for producing organisms with pesticide resistance.

XX PS Example 6; SEQ ID NO 50; 178pp; English.

XX CC This sequence represents a delta-endotoxin crystal protein. This protein
XX CC was included in the scope of the invention as a comparison to the delta-
XX CC endotoxins of the invention. Some of the delta-endotoxin coding sequences
XX CC of the invention have alternative start codons, producing more than one
XX CC protein from a single open reading frame. The nucleic acid sequences of
XX CC the invention are useful in DNA constructs or expression cassettes for
XX CC transformation and expression in plants and bacteria. The nucleic acids
XX CC and corresponding polypeptides are useful for killing lepidopteran or
XX CC coleopteran pests. Compositions containing the delta-endotoxins of the
XX CC invention, and methods for their production, are useful for the
XX CC production of organisms with pesticide resistance, specifically bacteria
XX CC and plants. These organisms are useful for generating altered or improved
XX CC delta-endotoxin or delta-endotoxin-associated proteins that have
XX CC pesticidal activity, or for detecting the presence of delta-endotoxin or
XX CC delta-endotoxin-associated proteins or nucleic acids in products or
XX CC organisms.

XX SQ Sequence 558 AA;

Query Match 20.6%; Score 62.5; DB 8; Length 558;
Best Local Similarity 27.7%; Pred. No. 50;
Matches 13; Conservative 8; Mismatches 17; Indels 9; Gaps 2;

OY 9 VHCTKNS-----LWGKVR-LKMKYDQNTTVMGRLLQDILLGWAT 46

Db 26 IECMSNEHSSKEBMLWDEVKQAKQLSWSRNLLYNGDFEDVSNMGKT 72

RESULT 5
AAW11633
ID AAW11633 standard; protein; 42 AA.

XX AC AAW11633;
XX DT 25-MAR-2003 (revised)
XX DT 05-JAN-1998 (first entry)

XX DE Lactobacillus plantarum bacteriocin.

XX KW bacteriocin; Lactobacillus plantarum BN85; DSM 9296; soft cheese;
XX KW inhibitor; growth; ripening; washed-rind; Munster; L. monocytogenes;
XX KW L. innocua; L. seeligeri; L. ivanovii; Listeria.

XX OS Lactobacillus plantarum.

XX PH Key Location/Qualifiers
XX FT Disulfide-bond 9..24
XX FT Modified-site 42
XX FT /note= "Cys-NH2"

XX PN DE29616369-UI.

XX PD 23-JAN-1997.

XX PF 20-SEP-1996; 96DE-02016369.

XX PR 20-SEP-1996; 96DE-02016369.

XX PA (FROM-) FROMAGERIE SIFFERT FRERES SA.

XX DR WPI; 1997-088733/09.

XX PT Lactobacillus strain that inhibits Listeria - for use in cheese prodn.

XX PS Example 4; Page 32; 53pp; German.

XX CC This is bacteriocin from a Lactobacillus strain [L. plantarum BN85 (DSM
XX CC 9296)], which is derived from soft cheese. A partial fragment derived by
XX CC Edman degradation (AAW11632) is capable of inhibiting growth of Listeria
XX CC bacteria in cheese at the start of ripening. The new Lactobacillus strain
XX CC is useful for prodn. of Listeria-free cheese, pref. washed-rind cheese,
XX CC esp. Munster cheese. Strain DSM 9296 is added to the cheese during
XX CC ripening, e.g. by spraying the cheese with a soln. contg. 104-106 cfu/ml
XX CC of Lactobacillus. DSM 9296 is capable of inhibiting Listeria
XX CC monocytogenes 1/2a, 1/2b, 1/2c, 3a, 3b, 4b, 4d and V7; Listeria innocua
XX CC 6a and 6b; Listeria seeligeri; and Listeria ivanovii. (Updated on 25-MAR-
XX CC 2003 to correct PR field.)

XX SQ Sequence 42 AA;

Query Match 20.4%; Score 62; DB 2; Length 42;
Best Local Similarity 39.1%; Pred. No. 2.6;
Matches 18; Conservative 6; Mismatches 8; Indels 14; Gaps 4;

OY 3 YYGTNGVHCTKNSL-WGKVRLLKMKYDQNTTVMGRLLQDILLGWATG 47
DB 2 YYG-NGVTCGKHSDWGKA-----TTCI--INNGAWAWATG 34

RESULT 6

AAW66446
ID AAW66446 standard; peptide; 37 AA.

XX AC AAW66446;

XX DT 12-JAN-1999 (first entry)

```

DE XX Cationic peptide leukocin A-val 187.
KW XX Indolicidin analogue; resistance; cationic peptide; antibiotic;
KW XX bacterial infection; tolerance; antibacterial; microorganism; bacteria;
KW XX fungus; parasite; virus.
XX OS Leuconostoc gelidum.
XX XX WO9840401-A2.
XX XX 17-SEP-1998.
XX XX
XX PF 10-MAR-1998; 98WO-CA000190.
XX PR 10-MAR-1997; 97US-0040649P.
XX PR 20-AUG-1997; 97US-00915314.
XX PR 26-SEP-1997; 97US-0060099P.
XX PR 25-FEB-1998; 98US-00030619.
XX XX
XX PA (MICR-) MICROLOGIX BIOTECH INC.
XX XX
XX PI Fraser JR, West MHP, Mcnicol PU;
XX XX
XX DR WPI; 1998-520800/44.
XX XX
XX PT New indolicidin peptide analogues - useful for, e.g. enhancing activity
XX PT of antibiotic or overcoming tolerance, acquired resistance or inherent
XX PT resistance of microorganisms.
XX PS Disclosure; Page 11; 105pp; English.
XX XX
XX CC AAW66393 to AAW66469 represent native cationic peptides from the present
XX CC invention. The present invention describes compositions and methods for
XX CC treating infection, especially bacterial infections. The compositions and
XX CC methods use cationic peptides in combination with an antibiotic agent
XX CC which are then administered to a patient to enhance the activity of the
XX CC antibiotic agent, to overcome: (a) tolerance; (b) acquired resistance;
XX CC and (c) inherent resistance. The combinations of antibiotics and cationic
XX CC peptides can provide synergistic activity against a microorganism that is
XX CC tolerant, inherently resistant, or has acquired resistance to an
XX CC antibiotic agent. They can be used for killing e.g. bacteria, fungi,
XX CC parasites and viruses
XX XX
XX SQ Sequence 37 AA;

Query Match 19.7%; Score 60; DB 2; Length 37;
Best Local Similarity 48.4%; Pred. No. 4;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;

Qy 3 YYGTNGVHCTKNSL---WGKV-----RLKN 24
   ||| ||||| |||
Db 2 YYG-NGVHCTKSGSVNWGEAFSAGVHRLAN 31

RESULT 7
AAY91745
ID AAY91745 standard; peptide; 37 AA.
XX AC
XX XX
XX XX AAY91745;
XX XX
XX DT 06-JUN-2000 (first entry)
XX XX
XX DE Cationic peptide Leukocin A-val 187 amino acid sequence.
XX XX
XX KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
XX KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
XX KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
XX KW multidrug resistance.
XX OS Unidentified.
XX XX
XX PN WO9965506-A2.
XX XX

Query Match 19.7%; Score 60; DB 3; Length 37;
Best Local Similarity 48.4%; Pred. No. 4;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;

Qy 3 YYGTNGVHCTKNSL---WGKV-----RLKN 24
   ||| ||||| |||
Db 2 YYG-NGVHCTKSGSVNWGEAFSAGVHRLAN 31

RESULT 8
AAR56481
ID AAR56481 standard; protein; 300 AA.
XX AC
XX XX
XX XX AAR56481;
XX XX
XX DT 25-MAR-2003 (revised)
XX DT 22-FEB-1995 (first entry)
XX XX
XX DE CD38.
XX XX
XX KW CD38; therapeutic; diagnostic.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO9417184-A1.
XX XX
XX PD 04-AUG-1994.
XX XX
XX PF 27-JAN-1994; 94WO-US000517.
XX XX
XX PR 29-JAN-1993; 93US-00010905.
XX XX
XX PA (SCHE ) SCHERING CORP.
XX PA (STRD ) UNIV LELAND STANFORD JUNIOR.
XX XX
XX PI Parkhouse RME, Santos-Argumedo L, Grimaldi JC, Bazan JF, Heath A;
XX PI Howard MC, Goodnow CC;
XX XX
XX DR WPI; 1994-264098/32.
XX DR N-PSDB; AAQ70644.
XX XX
XX PT Modulation of responses of lymphocytes - using antibody to CD38, a
XX PT soluble CD38 fragment or a modulator of ADP-ribosyl cyclase or cyclic ADP
XX PT -ribosyl hydrolase.

```

XX PS Disclosure; Page 46; 54pp; English.

XX CC The protein is a human CD38 molecule which is used to modulate the

XX CC physiological response of a lymphocyte, which is useful for the

XX CC stimulation/inhibition of lymphocyte growth or differentiation,

XX CC particularly for the establishment of antigen tolerance. (Updated on 25-

XX CC MAR-2003 to correct PN field.)

XX SQ Sequence 300 AA;

Query Match 19.7%; Score 60; DB 2; Length 300;

Best Local Similarity 35.0%; Pred. No. 49;

Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

Qy 5 GTNGVHCTKNSLWGKVRKNNKYDQNTTYMGRLODILGW 44

Db 113 GTQVPCNKILLWSRIKDLAHQFTQVQRDMFTLEDTLGY 152

RESULT 9

AAW37928

ID AAW37928 standard; protein; 300 AA.

XX AC AAW37928;

XX DT 09-SEP-1998 (first entry)

XX DE Amino acid sequence of CD38.

XX CD38; CD38 peptide fragment; anti-CD38 autoantibody; autoimmune disease;

XX KW insulin dependent diabetes mellitus; type II diabetes.

XX OS Mammalia.

XX PN WO9816245-A1.

XX PD 23-APR-1998.

XX PF 11-APR-1997; 97WO-JP001259.

XX PR 15-OCT-1996; 96JP-00272537.

XX PA (SHIO) SHIONOGI & CO LTD.

XX PI Taminato T;

XX WPI; 1998-251054/22.

XX N-PSDB; AAV29155.

XX Assaying anti-CD38 auto-antibody - useful for detecting auto-immune

XX disease, e.g. type II diabetes.

XX Disclosure; Page 8-10; 17pp; Japanese.

XX This is the amino acid sequence of the CD38 protein, of which a fragment

XX is used in the method of the invention to create the anti-CD38

XX autoantibody. The method is used for detecting autoimmune disease, e.g.

XX insulin dependent diabetes mellitus or type II diabetes

XX SQ Sequence 300 AA;

Query Match 19.7%; Score 60; DB 2; Length 300;

Best Local Similarity 35.0%; Pred. No. 49;

Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

Qy 5 GTNGVHCTKNSLWGKVRKNNKYDQNTTYMGRLODILGW 44

Db 113 GTQVPCNKILLWSRIKDLAHQFTQVQRDMFTLEDTLGY 152

RESULT 10

AAB69069

ID AAB69069 standard; protein; 300 AA.

XX AC AAB69069;

XX DT 19-APR-2001 (first entry)

XX DE Human CD38 protein sequence SEQ ID NO:2.

XX KW Human; CD38; diabetes mellitus; detection; cyclic ADP-ribose; CADPR.

XX OS Homo sapiens.

XX JP2000316578-A.

XX PD 21-NOV-2000.

XX PF 12-MAY-1999; 99JP-00131955.

XX PR 12-MAY-1999; 99JP-00131955.

XX PA (BMLB-) BML KK.

XX PA (KANE/) KANETSUKA A.

XX PA (OKAM/) OKAMOTO H.

XX DR WPI; 2001-128255/14.

XX DR N-PSDB; AAF32487.

XX Detecting onset of diabetes mellitus comprises detecting specific gene

XX mutations in the CD38 gene.

XX Example; Page 11-12; 19pp; Japanese.

XX The present invention describes a method using a mutation in the CD38

XX gene (involved in the production of cyclic ADP-ribose (CADPR)), to detect

XX the onset of diabetes mellitus. The method is useful for detecting the

XX onset of diabetes mellitus. The present sequence represents human CD38,

XX which is used in an example from the present invention

XX SQ Sequence 300 AA;

Query Match 19.7%; Score 60; DB 4; Length 300;

Best Local Similarity 35.0%; Pred. No. 49;

Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

Qy 5 GTNGVHCTKNSLWGKVRKNNKYDQNTTYMGRLODILGW 44

Db 113 GTQVPCNKILLWSRIKDLAHQFTQVQRDMFTLEDTLGY 152

RESULT 11

ABG61817

ID ABG61817 standard; protein; 300 AA.

XX AC ABG61817;

XX DT 15-AUG-2002 (first entry)

XX DE Prostate cancer-associated protein #18.

XX KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.

XX OS Mammalia.

XX WO200230268-A2.

XX PD 18-APR-2002.

XX PF 12-OCT-2001; 2001WO-US032045.

XX PR 13-OCT-2000; 2000US-00687576.

XX PR 08-DEC-2000; 2000US-00733288.

XX PR 08-DEC-2000; 2000US-00733742.

XX PR 24-JAN-2001; 2001US-0263957P.

```

PR 16-MAR-2001; 2001US-0276791P.
PR 16-MAR-2001; 2001US-0276888P.
PR 06-APR-2001; 2001US-0281922P.
PR 24-APR-2001; 2001US-0286214P.
PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0288589P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX
XX WPI; 2002-471335/50.
XX N-PSDB; ABK92132.
XX
PT Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX
XX Claim 27; Page 314; 436pp; English.
XX
XX The present invention relates to methods of detecting a prostate cancer-
XX associated transcript in a cell from a patient. The method comprises
XX contacting a biological sample from the patient with prostate cancer-
XX associated polynucleotides (designated PC genes) that selectively
XX hybridise to a sequence that is at least 80% identical to them. The
XX prostate cancer-associated polynucleotide sequences are differentially
XX expressed in prostate tumour tissue or in prostate cancer and are derived
XX from the tissues of various organisms such as humans or other mammals
XX (e.g. mice, sheep and dogs). The methods of the invention are useful for
XX diagnosing and treating prostate cancer in mammals. The prostate cancer-
XX associated genes are useful for diagnosing or treating prostate cancer,
XX as well as for identifying modulators of prostate cancer or agents that
XX inhibit prostate cancer. The nucleic acid sequences are particularly
XX useful in gene therapy, as a vaccine or in antisense applications.
XX ABG61800-ABG61944 represent prostate cancer-associated proteins
XX
XX SQ Sequence 300 AA;
XX
XX Query Match 19.7%; Score 60; DB 5; Length 300;
XX Best Local Similarity 35.0%; Pred. No. 49;
XX Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
XX
Oy 5 GTNGVCHTNSLGMGKVRLLKMKYDQNTYMGRLQDILLGW 44
Db ||| ||| ||| ::| ||| ||| |||
113 GTQVPCNKILLWSRIKDLAQHTQVQVRDMFTLEDTLIGY 152

RESULT 12
ABU09711
ID ABU09711 standard; protein; 300 AA.
XX
XX AC ABU09711;
XX
XX DT 03-JUL-2003 (first entry)
XX
XX DE CD38 protein.
XX
XX KW CD38; diabetic onset; diabetes; lymphocyte surface marker.
XX
XX OS Hominidae.
XX
XX FH Key Location/Qualifiers
XX Misc-difference 264
XX FT /note= "Wild type Ser substituted by Leu"
XX
XX PN US2003027134-A1.
XX
XX PD 06-FEB-2003.
XX
XX PF 31-JAN-2001; 2001US-00773307.
XX
XX PR 31-JAN-2001; 2001US-00773307.
XX
XX PA (EGAS/) EGASHIRA T.
XX PA (NAGA/) NAGANO M.
XX PA (SAGE/) SAGEHASHI Y.
XX PF 31-JAN-2001; 2001US-00773307.
XX PR 31-JAN-2001; 2001US-00773307.
XX PA (EGAS/) EGASHIRA T.
XX PA (NAGA/) NAGANO M.
XX PA (SAGE/) SAGEHASHI Y.
XX
XX EGashira T, Nagano M, Sagehashi Y, Matsui K, Hattori H;
XX (SAGE/) SAGEHASHI Y.
XX

```

```

PA (MATS/) MATSUI K.
PA (HATT/) HATTORI H.
PA (KANA/) KANATSUKA A.
PA (TAKA/) TAKASAWA S.
PA (OKAM/) OKAMOTO H.
XX
XX EGashira T, Nagano M, Sagehashi Y, Matsui K, Hattori H;
XX PI Kanatsuka A, Takasawa S, Okamoto H;
XX
XX WPI; 2003-417248/39.
XX N-PSDB; ACA60951.
XX
XX DR Detecting risk factor for onset of diabetes in an individual, involves
XX detecting genetic abnormality of the gene CD38.
XX
XX PS Disclosure; Page 10; 29pp; English.
XX
XX CC The invention describes a method of detecting a risk factor for diabetic
XX onset in an individual. The method involves detecting genetic abnormality
XX of the gene CD38 which is a human lymphocyte surface marker and is a risk
XX gene for onset of diabetes. The sites of abnormality in CD38 gene include
XX CC a site encoding Arg at residue 140, a site encoding Ser at residue 264,
XX CC or guanine at nucleotide position -28 in intron 7. This is the amino acid
XX sequence of wild type CD38 protein
XX
XX SQ Sequence 300 AA;
XX
XX Query Match 19.7%; Score 60; DB 6; Length 300;
XX Best Local Similarity 35.0%; Pred. No. 49;
XX Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
XX
Oy 5 GTNGVCHTNSLGMGKVRLLKMKYDQNTYMGRLQDILLGW 44
Db ||| ||| ||| ::| ||| ||| |||
113 GTQVPCNKILLWSRIKDLAQHTQVQVRDMFTLEDTLIGY 152

RESULT 13
ABU09714
ID ABU09714 standard; protein; 300 AA.
XX
XX AC ABU09714;
XX
XX DT 03-JUL-2003 (first entry)
XX
XX DE CD38 protein S264L mutant.
XX
XX KW CD38; diabetic onset; diabetes; lymphocyte surface marker; mutant;
XX mutein.
XX
XX OS Hominidae.
XX
XX FH Key Location/Qualifiers
XX Misc-difference 264
XX FT /note= "Wild type Ser substituted by Leu"
XX
XX PN US2003027134-A1.
XX
XX PD 06-FEB-2003.
XX
XX PF 31-JAN-2001; 2001US-00773307.
XX
XX PR 31-JAN-2001; 2001US-00773307.
XX
XX PA (EGAS/) EGASHIRA T.
XX PA (NAGA/) NAGANO M.
XX PA (SAGE/) SAGEHASHI Y.
XX PA (MATS/) MATSUI K.
XX PA (HATT/) HATTORI H.
XX PA (KANA/) KANATSUKA A.
XX PA (TAKA/) TAKASAWA S.
XX PA (OKAM/) OKAMOTO H.
XX
XX EGashira T, Nagano M, Sagehashi Y, Matsui K, Hattori H;
XX (SAGE/) SAGEHASHI Y.
XX

```

XX
SD

CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 300 AA;

Query Match 19.7%; Score 60; DB 7; Length 300;
Best Local Similarity 35.0%; Pred. No. 49;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
Qy 5 GTNGVHCTKNSLWGKVRKNNKMYDQNTTYMGRQLQDILLGW 44
Db 113 GTQVPCNKILLWSRIKDLAHQFTQVQDRDMFTLEDTLGY 152

Search completed: May 2, 2006, 17:26:33
Job time : 191 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 17:26:51 ; Search time 38 Seconds
(without alignments)
136.729 Million cell updates/sec

Title: US-10-644-927-1
Perfect score: 304
Sequence: 1 KTYVGTNGVHCTKNSLWGVK.....GRLQDILLGWATGAFGKTFH 54

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| 1 | 65.5 | 21.5 | 471 | 2 S11899 | glutamate-ammonia |
| 2 | 65.5 | 21.5 | 471 | 2 A99180 | hypothetical prote |
| 3 | 64 | 21.1 | 61 | 1 S38508 | sakacin P precursor |
| 4 | 63 | 20.7 | 590 | 2 AF2555 | hypothetical prote |
| 5 | 62.5 | 20.6 | 446 | 2 A43995 | glutamate-ammonia |
| 6 | 60 | 19.7 | 61 | 2 A41657 | leucocin A-UAL 187 |
| 7 | 60 | 19.7 | 61 | 2 S52208 | mesentericin Y105 |
| 8 | 60 | 19.7 | 300 | 2 A43521 | lymphocyte surface |
| 9 | 59.5 | 19.6 | 286 | 2 P84954 | spermidine synthas |
| 10 | 59 | 19.4 | 62 | 1 A48941 | pediocin PA-1 prec |
| 11 | 59 | 19.4 | 399 | 2 T20455 | hypothetical prote |
| 12 | 58 | 19.1 | 3036 | 2 T18995 | hypothetical prote |
| 13 | 57.5 | 18.9 | 510 | 2 T50021 | inositol-3-phospha |
| 14 | 56.5 | 18.6 | 517 | 1 G18PT4 | gene 12 protein - |
| 15 | 55.5 | 18.3 | 132 | 2 PC2131 | hepatocyte growth |
| 16 | 55.5 | 18.3 | 934 | 2 B28838 | parasporal crystal |
| 17 | 55.5 | 18.3 | 1332 | 2 F69732 | PBSX prophage ORF |
| 18 | 55 | 18.1 | 201 | 2 AD3302 | 31K outer-membrane |
| 19 | 55 | 18.1 | 346 | 2 B20243 | aminomethyltransfe |
| 20 | 55 | 18.1 | 369 | 2 T24205 | hypothetical prote |
| 21 | 55 | 18.1 | 516 | 2 G84442 | probable nucleosid |
| 22 | 55 | 18.1 | 799 | 1 TVRTTB | nerve growth facto |
| 23 | 54.5 | 17.9 | 319 | 2 T15463 | hypothetical prote |
| 24 | 54.5 | 17.9 | 322 | 2 D86760 | phosphate starvati |
| 25 | 54.5 | 17.9 | 374 | 2 F97257 | glycosyltransferas |
| 26 | 54.5 | 17.9 | 383 | 2 B96806 | hypothetical prote |
| 27 | 54.5 | 17.9 | 494 | 1 B41141 | 1-aminocyclopropan |
| 28 | 54 | 17.8 | 262 | 1 TLBPM1 | tail fiber protein |
| 29 | 54 | 17.8 | 269 | 2 S51815 | hypothetical prote |

ALIGNMENTS

RESULT 1

S11899
glutamate-ammonia ligase (EC 6.3.1.2) - Sulfolobus solfataricus
N:Alternate names: glutamine synthetase
C:Species: Sulfolobus solfataricus
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S11899
R:Sanangelantoni, A.M.; Barbarini, D.; di Pasquale, G.; Cammarano, P.; Tiboni, O.
Mol. Gen. Genet. 221, 187-194, 1990
A:Title: Cloning and nucleotide sequence of an archaeobacterial glutamine synthetase gene
A:Reference number: S11899; MUID:90318316; PMID:1973523
A:Accession: S11899
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-471 <SAN>
A:Cross-references: UNIPROT:P23794; UNIPARC:UPI0000170275; GB:X53263; NID:G296864; PIDN
C:Superfamily: glutamate-ammonia ligase
C:Keywords: ligase

Query Match 21.5%; Score 65.5; DB 2; Length 471;
Best Local Similarity 48.6%; Pred. No. 2;
Matches 17; Conservative 2; Mismatches 13; Indels 3; Gaps 2;

QY 1 KTYVGTN--GVHCTKNSLWGVKVRUKMKYDQNTTY 33
DB 255 KPIYGDNGTGMH--THLSLWTKDGGKQKMLYDPNDEY 288

RESULT 2

A99180
hypothetical protein glna-1 [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: A99180
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, J.
arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: A99180
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-471 <KUR>
A:Cross-references: UNIPROT:P23794; UNIPARC:UPI000012B704; GB:AE006641; NID:gl3813510;
C:Genetics:
C:Superfamily: glutamate-ammonia ligase

Query Match 21.5%; Score 65.5; DB 2; Length 471;
Best Local Similarity 48.6%; Pred. No. 2;
Matches 17; Conservative 2; Mismatches 13; Indels 3; Gaps 2;

Qy 1 KTYVGTN--GVHCTKNSLWGVRLKNMKYDQNTTY 33
| | | | | : | | | | | : | | | | |
Db 255 KPIYGDNGTGMH--THLSLWTKDGKKNLWYDPNDEY 288
| | | | | : | | | | | : | | | | |
RESULT 3
S38508
sakacin P precursor - Lactobacillus sake (strains Lb674 and LTH673)
A;Variety: Lactobacillus sake
C;Species: Lactobacillus sake
C;Date: 25-Dec-1994 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S57911; S57915; S43689; S38508
R;Huehne, K.; Holck, A.; Axelsson, L.; Axelsson, L.
submitted to the EMBL Data Library, February 1995
A;Description: Purification and cloning of sakacin 674, a bacteriocin from Lactobacillus
A;Reference number: S57910
A;Accession: S57911
A;Molecule type: DNA
A;Residues: 1-61 <HUE>
A;Cross-references: UNIPROT:P35618; UNIPARC:UPI00000013AD; EMBL:Z48542; NID:G695615; PID
A;Experimental source: strain Lb674
R;Holck, A.L.; Axelsson, L.; Huehne, K.; Kroeckel, L.
FEMS Microbiol. Lett. 115, 143-150, 1994
A;Title: Purification and cloning of sakacin 674, a bacteriocin from Lactobacillus sake
A;Reference number: S57915; MUID:94186010; PMID:8138128
A;Accession: S57915
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-61 <HOL>
A;Cross-references: UNIPARC:UPI00000013AD; EMBL:Z25816; NID:9414124; PIDN:CAA81064.1; PI
A;Experimental source: strain Lb674
R;Tichaczek, P.S.; Vogel, R.F.; Hammes, W.P.
Microbiology 140, 361-367, 1994
A;Title: Cloning and sequencing of sakP encoding sakacin P, the bacteriocin produced by
A;Reference number: S43688; MUID:94236240; PMID:8180701
A;Accession: S43689
A;Molecule type: DNA
A;Residues: 1-61 <TIC>
A;Cross-references: UNIPARC:UPI00000013AD; EMBL:X75081; NID:9475950; PIDN:CAA52974.1; PI
A;Experimental source: strain LTH 673
C;Genetics:
A;Gene: sakP; sppA
C;Superfamily: carnobacteriocin precursor
C;Keywords: antibacterial; bacteriocin
F;1-18/Domain: propeptide #status predicted <PRO>
F;19-61/Product: sakacin P #status predicted <MAT>
F;27-32/Disulfide bonds: #status predicted

Query Match 21.1%; Score 64; DB 1; Length 61;
Best Local Similarity 36.7%; Pred. No. 0.31;
Matches 18; Conservative 4; Mismatches 9; Indels 18; Gaps 4;
Qy 3 YYGTNGVHCTKNSL---WGKVRKLNKMYDQNTTYMGRLL-QDILLGWATG 47
| | | | | | | | | | | | | | | : | | | | |
Db 20 YYG-NGVHCTKGKHSCTVDWG-----TAIGNIGNNAANWATG 54
| | | | | | | | | | | | | | | : | | | | |
RESULT 4
AF2555
hypoetical protein all18037 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120ga
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AF2555
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Saamamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AF2555
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-590 <KUR>
A;Cross-references: UNIPROT:Q8YK80; UNIPARC:UPI000000CEFF0; GB:AP003603; PIDN:BAB77367.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all18037
A;Genome: plasmid

Query Match 20.7%; Score 63; DB 2; Length 590;
Best Local Similarity 35.3%; Pred. No. 5.4;
Matches 18; Conservative 8; Mismatches 23; Indels 2; Gaps 1;
Qy 3 YYGTNGVHCTKNSLWGVRLKNMKYDQNTTYMGRLLQDILLGWATGAFGKTF 53
| | | | | : | | | | | : | | | | |
Db 88 YIGT--PRGTFQVGVGNKRITNIPEDKNRLYLDPVQRGILVSGSGSGKTF 136
| | | | | : | | | | | : | | | | |
RESULT 5
A43995
glutamate-ammonia ligase (EC 6.3.1.2) - Methanococcus voltae
N;Alternate names: glutamine synthetase
C;Species: Methanococcus voltae
C;Date: 11-Feb-1993 #sequence_revision 11-Feb-1993 #text_change 09-Jul-2004
C;Accession: A43995; S13504
R;Possot, O.; Sibold, L.; Aubert, J.P.
Res. Microbiol. 140, 355-371, 1989
A;Title: Nucleotide sequence and expression of the glutamine synthetase structural gene,
A;Reference number: A43995; MUID:90139872; PMID:2575777
A;Accession: A43995
A;Molecule type: DNA
A;Residues: 1-446 <POS>
A;Cross-references: UNIPROT:P21154; UNIPARC:UPI000012B6EC; GB:X53509; NID:944711; PIDN:C
C;Genetics:
A;Gene: glnA
C;Superfamily: glutamate-ammonia ligase
C;Keywords: ligase

Query Match 20.6%; Score 62.5; DB 2; Length 446;
Best Local Similarity 33.9%; Pred. No. 4.5;
Matches 19; Conservative 9; Mismatches 17; Indels 11; Gaps 5;
Qy 1 KTYVGTN--GVHCTKNSLW--GKVRKLNKMYDQNTTYMGRLLQDILLGWATGAFGKT 52
| | | | | : | | | | | : | | | | |
Db 237 KPFFGVNGSGMHCNQ-SIMLDGKPSF----YDENNAH--QLSDBICLSYIGGILEHT 285
| | | | | : | | | | | : | | | | |
RESULT 6
A41657
leucocin A-UAL 187 precursor - Leuconostoc gelidium plasmid pLG7.6
C;Species: Leuconostoc gelidium
C;Date: 13-May-1992 #sequence_revision 13-May-1992 #text_change 09-Jul-2004
C;Accession: A41657
R;Hastings, J.W.; Sailer, M.; Johnson, K.; Roy, K.L.; Vederas, J.C.; Stiles, M.E.
J. Bacteriol. 173, 7491-7500, 1991
A;Title: Characterization of leucocin A-UAL 187 and cloning of the bacteriocin gene from
A;Reference number: A41657; MUID:92041660; PMID:1840587
A;Accession: A41657
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-61 <HAS>
A;Cross-references: UNIPROT:P34034; UNIPARC:UPI000012E281; GB:M64371; NID:9149637; PIDN:J
C;Genetics:
A;Genome: plasmid
C;Superfamily: carnobacteriocin precursor
C;Keywords: bacteriocin

Query Match 19.7%; Score 60; DB 2; Length 61;
Best Local Similarity 48.4%; Pred. No. 1;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;
Qy 3 YYGTNGVHCTKNSL---WGKV-----RLKN 24
| | | | | | | | | | | | | | | : | | | | |
Db 26 YYG-NGVHCTKSGCVNMGEAFSAGVHRLAN 55
| | | | | | | | | | | | | | | : | | | | |

Search completed: May 2, 2006, 17:31:16
Job time : 41 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 17:23:40 ; Search time 234 Seconds
(without alignment)
162.814 Million cell updates/sec

Title: US-10-644-927-1

Perfect score: 304

Sequence: 1 KTYGTNGVHCTKSLGWKV.....GRLQDILLGWATGAFGKTFH 54

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------|--------------------|
| 1 | 195 | 64.1 | 81 | 2 Q48496_LACAC | Q48496 lactobacill |
| 2 | 65.5 | 21.5 | 471 | 1 GLNA_SULSO | P23794 sulfolobus |
| 3 | 65 | 21.4 | 301 | 1 CD38_MACFA | Q5van0 macaca fabc |
| 4 | 65 | 21.4 | 325 | 2 Q5XQ46_9PEZI | Q5xq46 cryptovalia |
| 5 | 64.5 | 21.2 | 460 | 2 Q4NH89_9MICC | Q4nh89 arthrobacte |
| 6 | 64.5 | 21.2 | 1071 | 2 Q7SCA6_ASHGO | Q7sca6 ashbya gos |
| 7 | 64 | 21.1 | 61 | 1 SAKP_LACSK | P35618 lactobacill |
| 8 | 64 | 21.1 | 61 | 2 Q7B5F5_LACSK | Q7b5f5 lactobacill |
| 9 | 64 | 21.1 | 714 | 2 Q51P75_MAGGR | Q51p75 magnaporth |
| 10 | 63 | 20.7 | 590 | 2 Q8YK80_ANAEP | Q8yk80 anabaena sp |
| 11 | 63 | 20.7 | 1582 | 2 Q8RIM1_FUSNN | Q8rim1 fusobacteri |
| 12 | 63 | 20.7 | 1630 | 2 Q8RHH7_FUSNN | Q8rhh7 fusobacteri |
| 13 | 62.5 | 20.6 | 446 | 1 GLNA_METVO | P21154 methanococ |
| 14 | 62.5 | 20.6 | 558 | 2 Q8VM63_BACTA | Q8vm63 bacillus th |
| 15 | 61.5 | 20.2 | 469 | 2 Q9PZ03_GVXN | Q9pz03 xestia c-ni |
| 16 | 61.5 | 20.2 | 524 | 2 Q7QBD6_ANOGA | Q7qed6 anopheles g |
| 17 | 61.5 | 20.2 | 547 | 2 Q8A1J6_BACTN | Q8a1j6 bacteroides |
| 18 | 61.5 | 20.2 | 853 | 2 Q9DL93_9HIV1 | Q9dl93 human immun |
| 19 | 61.5 | 20.2 | 858 | 2 Q9DL79_9HIV1 | Q9dl79 human immun |
| 20 | 61.5 | 20.2 | 1474 | 2 Q4Q2Y5_LEIMA | Q4q2y5 leishmania |
| 21 | 61 | 20.1 | 61 | 2 Q4U1B4_LACCU | Q4uib4 lactobacill |
| 22 | 61 | 20.1 | 684 | 2 Q5WI64_BACSK | Q5wi64 bacillus cl |
| 23 | 61 | 20.1 | 1524 | 2 Q4UIZ1_THEAN | Q4uiz1 theileria a |
| 24 | 60 | 19.7 | 61 | 1 LCCA_LEUGE | P34034 leuconostoc |
| 25 | 60 | 19.7 | 61 | 1 LCCB_LEUCA | Q53446 leuconostoc |
| 26 | 60 | 19.7 | 61 | 1 MTCV_LEUME | P38577 leuconostoc |
| 27 | 60 | 19.7 | 61 | 2 Q791V9_LEUME | Q791v9 leuconostoc |
| 28 | 60 | 19.7 | 300 | 1 CD38_HUMAN | P28907 homo sapien |
| 29 | 60 | 19.7 | 498 | 2 Q622U2_CAEBR | Q622u2 caenorhabdi |
| 30 | 60 | 19.7 | 1143 | 2 Q948Y6_VOLCA | Q948y6 volvox cart |
| 31 | 60 | 19.7 | 1794 | 2 Q8RHH1_FUSNN | Q8rhh1 fusobacteri |

RESULT 1

Q48496.LACAC
ID Q48496_LACAC PRELIMINARY; PRT; 81 AA.
AC Q48496;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Acidocin A precursor.
GN Name=acdA;
OS Lactobacillus acidophilus.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1579;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TK9201;
RX MEDLINE=95314239; PubMed=7793908;
RA Kanatani K., Oshimura M., Sano K.;
RT "Isolation and characterization of acidocin A and cloning of the
RT bacteriocin gene from Lactobacillus acidophilus.";
RL Appl. Environ. Microbiol. 61:1061-1067(1995).
DR EMBL; D37881; BAA07120.1; -; Genomic_DNA.
DR InterPro; IPR010133; Bacteriocin_sig.
DR TIGRFAMs; TIGR01847; bacteriocin_sig; 1.
KW Signal.
FT SIGNAL. 1 23 Potential.
FT CHAIN 24 81 mature acidocin A.
SQ SEQUENCE 81 AA; 8946 MW; 89698AA296F7819D CRC64;
Query Match 64.1%; Score 195; DB 2; Length 81;
Best Local Similarity 64.6%; Pred. No. 2.3e-17;
Matches 42; Conservative 1; Mismatches 4; Indels 18; Gaps 3;

QY 1 KTYGTNGVHCTKSLGWKVRLKNMKYDONTTYMGRLL-----QD---ILLGWATGAF 49
Db 24 KTYGTNGVHCTKSLGWKVRLKNV-----IPGTLCKRQSLPIKQDLKILLGWATGAF 76
QY 50 GKTFH 54
Db 77 GKTFH 81

RESULT 2

GLNA_SULSO
ID GLNA_SULSO STANDARD; PRT; 471 AA.
AC P23794;
DT 01-NOV-1991 (Rel. 20, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase) (GS).
GN Name=glina; Synonyms=glina-1; OrderedLocusNames=SSO0366;
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.

```

OX NCBI_TaxID=2287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=90318316; PubMed=1973523; DOI=10.1007/BF00261719;
RA Sanangelantoni A.M., Barbarini D., di Pasquale G., Cammarano P.,
RA Tiboni O.;
RT "Cloning and nucleotide sequence of an archaeobacterial glutamine
RT synthetase gene: phylogenetic implications.";
RL Mol. Gen. Genet. 221:187-194(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 35092 / DSM 1617 / P2; DOI=10.1073/pnas.141222098;
RX MEDLINE=21332296; PubMed=11427726; DOI=10.1073/pnas.141222098;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweys M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Eraso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -|- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +
CC L-glutamine.
CC -|- ENZYME REGULATION: The activity of this enzyme is controlled by
CC adenylation under conditions of abundant glutamine. The fully
CC adenylated enzyme complex is inactive (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: Belongs to the glutamine synthetase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X53263; CAA37353.1; -; Genomic DNA.
DR EMBL; X53263; CAA37352.1; -; Genomic DNA.
DR EMBL; AE006669; AAK40696.1; -; Genomic DNA.
DR PIR; A99180; A99180.
DR PIR; S11899; S11899.
DR HSP; P08201; ILGR.
DR InterPro; IPR008147; Gln_synth_beta.
DR InterPro; IPR008146; Gln_synth_C.
DR InterPro; IPR004809; GlnA.
DR Pfam; PF00120; Gln-synth_C; 1.
DR Pfam; PF03951; Gln-synth_N; 1.
DR ProDom; PD001057; Gln_synth_C; 1.
DR TIGRFAMs; TIGR00653; GlnA; 1.
DR PROSITE; PS00180; GlnA_1; 1.
DR PROSITE; PS00181; GlnA_ATP; 1.
DR Complete proteome; Ligase.
KW BINDING 400 400 AMP (covalent) (By similarity).
FT CONFLICT 162 162 S -> T (in Ref. 1).
SQ SEQUENCE 471 AA; 53348 MW; 1CC4239287B6C2CD CRC64;

Query Match 21.5%; Score 65.5; DB 1; Length 471;
Best Local Similarity 48.6%; Pred. No. 11;
Matches 17; Conservative 2; Mismatches 13; Indels 3; Gaps 2;

Qy 1 KTYYGTN--GVHCTKNSLWGKVRLLKMKYDQNTTY 33
| | | | | | | | | | | | | | | | | |
Db 255 KPIYGDNGTGMH--THLSLWTKDGKKNLWYDPNDEY 288

RESULT 3
CD38 MACFA STANDARD; PRT; 301 AA.
AC QSVANQ;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ADP-ribosyl cyclase 1 (EC 3.2.2.5) (Cyclic ADP-ribose hydrolase 1)

```

```

DE (cADPr hydrolase 1) (CD38 homolog).
GN Name=CD38;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RX PubMed=15383153; DOI=10.1186/1471-2172-5-21;
RA Ferrero E., Orciani M., Vacca P., Ortolan E., Crovella S., Titti F.,
RA Sacucci F., Malavasi F.;
RT "Characterization and phylogenetic epitope mapping of CD38 ADPR
RT cyclase in the cynomolgus macaque.";
RL BMC Immunol. 5:21-21(2004).
CC -|- FUNCTION: Synthesizes cyclic ADP-ribose, a second messenger for
CC glucose-induced insulin secretion. Also has cADPr hydrolase
CC activity. Also moonlights as a receptor in cells of the immune
CC system (By similarity).
CC -|- CATALYTIC ACTIVITY: NAD(+) + H(2)O = ADP-ribose + nicotinamide.
CC -|- ENZYME REGULATION: ATP inhibits the hydrolyzing activity (By
CC similarity).
CC -|- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).
CC -|- SIMILARITY: Belongs to the ADP-ribosyl cyclase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AY55148; AAT36330.1; -; mRNA.
DR InterPro; IPR003193; Rib_hydrolase.
DR PANTHER; PTHR10912; Rib_hydrolase; 1.
DR Pfam; PF02267; Rib_hydrolase; 1.
KW Glycoprotein; Hydrolase; NAD; Receptor; Signal-anchor; Transmembrane.
FT TOPO_DOM 1 21 Cytoplasmic (Potential).
FT TRANSMEM 22 43 Signal-anchor for type II membrane
FT protein (Potential).
FT TOPO_DOM 44 301 Extracellular (Potential).
FT ACT_SITE 120 120 By similarity.
FT ACT_SITE 202 202 By similarity.
FT CARBOHYD 101 101 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 121 121 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 210 210 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 220 220 N-linked (GlcNAc... ) (Potential).
FT DISULFID 68 83 By similarity.
FT DISULFID 100 181 By similarity.
FT DISULFID 161 174 By similarity.
FT DISULFID 255 276 By similarity.
FT DISULFID 288 297 By similarity.
SQ SEQUENCE 301 AA; 34422 MW; E659212B92616591 CRC64;

Query Match 21.4%; Score 65; DB 1; Length 301;
Best Local Similarity 35.0%; Pred. No. 8.1;
Matches 14; Conservative 8; Mismatches 18; Indels 0; Gaps 0;

Qy 5 GTGVHCTKNSLWGKVRLLKMKYDQNTTYMGRLODILLGW 44
| | | | | | | | | | | | | | | | | |
Db 114 GTQTPCNKTLMSRIKDLAHQFTQVORDMTLEDMLGY 153

RESULT 4
Q5XQ46_9PEZI
ID Q5XQ46_9PEZI PRELIMINARY; PRT; 325 AA.
AC Q5XQ46;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Family 10 xylanase (EC 3.2.1.8).
OS Cryptovalsa sp. BCC 7197.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Xylariomycetidae; Xylariales; Diatrypeaceae; Cryptovalsa.

```



```

OX NCB1_TaxID=295079;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BCC 7197;
RA Boonyapakorn K., Pootanakit K., Chantasingh D., Kirtikara K.,
RA Eurwilaichitr L.;
RT "Cloning and expression of xylanase 10 from Cryptovalsa sp. (BCC7197)
RT in Pichia pastoris.";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY741212; AAU89274.1; -; mRNA.
DR GO; GO:0031176; F:endo-1,4-beta-xylanase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001000; Glyco_hydro_10.
DR Pfam; PF00331; Glyco_hydro_10; 1.
DR PRINTS; PM00134; GLYHDLASE10.
DR SMART; SM00633; Glyco_10; 1.
DR KW Glycosidase; Hydrolase; Xylan degradation.
SQ SEQUENCE 325 AA; 35107 MW; 6E0930A7F8C6433C CRC64;

Query Match 21.4%; Score 65; DB 2; Length 325;
Best Local Similarity 32.8%; Pred. No. 8.8;
Matches 21; Conservative 6; Mismatches 15; Indels 22; Gaps 4;

QY 1 KTYTGTVGHVCHTKNSL-----WGKVRLEN-MKYDQNTYMGRLQ-----DILL 42
Db 36 KLYTGT-----CTDGLLTSGQSAIIQTNFGQVTPENSMKWDQTNSEKGFNLQAQADYLV 91

QY 43 GWAT 46
Db 92 DWAT 95

RESULT, 5
Q4NH89 9MICC PRELIMINARY; PRT; 460 AA.
AC Q4NH89
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Putative membrane protein precursor.
GN ORFNames=ArthDRAFT_2550;
OS Arthrobacter sp. FB24
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Micrococcaceae; Arthrobacter.
OX NCB1_TaxID=290399;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Izrasi S., Pitluck S., Richardson P.;
RA "Sequencing of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (PGF-ORNL);
RA Larimer F., Land M.;
RA "Annotation of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHG01000004; EAL97009.1; -; Genomic_DNA.
KW Signal.
FT SIGNAL 460 460 Potential.
SQ SEQUENCE 460 AA; 49805 MW; F5FB146AFB237113 CRC64;

Query Match 21.2%; Score 64.5; DB 2; Length 460;
Best Local Similarity 27.5%; Pred. No. 15;
Matches 19; Conservative 12; Mismatches 21; Indels 17; Gaps 3;

QY 1 KTYTGTVGHVCHTKNSLWGKVRLE-----KNMKYDQNTYMGRLQDILLG-----W 44

```

```

Db 158 RSTFGTSGPHATRNK-WGAVILLVNLGFWRETLAFGQINILMLGMAADLLARNQRW 216
QY 45 ATGAFGKTF 53
Db 217 NQGFPGRGF 225

RESULT 6
Q75CA6 ASHGO PRELIMINARY; PRT; 1071 AA.
AC Q75CA6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ACR006CP.
GN Name=ACR006C;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCB1_TaxID=33169;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 10895;
RC PubMed=15001715; DOI=10.1126/science.1095781;
RX Dietrich F.S., Voegelé S., Brachat S., Lerch A., Gates K., Steiner S.,
RA Mohr C., Poehlmann R., Luedi P., Choi S., Wing R.A., Flavier A.,
RA Gaffney T.D., Philippsen P.;
RA "The Ashbya gossypii genome as a tool for mapping the ancient
RT Saccharomyces cerevisiae genome.";
RL Science 304:304-307(2004).
DR EMBL; AE016816; AAS51233.1; -; Genomic_DNA.
DR AGD; ACR006C; -
KW Complete proteome.
SQ SEQUENCE 1071 AA; 123573 MW; 0816B52D60032663 CRC64;

Query Match 21.2%; Score 64.5; DB 2; Length 1071;
Best Local Similarity 31.7%; Pred. No. 39;
Matches 13; Conservative 8; Mismatches 11; Indels 9; Gaps 1;

QY 8 GVHCTKNSLWGKVRLENMKYDQNTYMGRLQDILLGWATGA 48
Db 702 GTAMAKNSLYGCLSVHNIKYEENAWYL-----WITSS 733

RESULT 7
SAKP LACSK STANDARD; PRT; 61 AA.
AC P35618; Q57121;
DT 01-JUN-1994 (Rel. 29, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Bacteriocin sakacin P precursor (Sakacin 674).
GN Name=sakP; Synonyms=sakR, appa;
OS Lactobacillus sakei.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCB1_TaxID=1599;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=LTH673;
RX MEDLINE=94236240; PubMed=8180701;
RA Tichaczek P.S., Vogel R.F., Hammes W.P.;
RT "Cloning and sequencing of sakP encoding sakacin P, the bacteriocin
RT produced by Lactobacillus sake LTH 673.";
RL Microbiology 140:361-367(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE, PROTEIN SEQUENCE OF 19-61, AND CHARACTERIZATION.
RC STRAIN=Lb674;
RX MEDLINE=94186010; PubMed=8138128; DOI=10.1016/0378-1097(94)90005-1;
RA Holck A.L., Axelsson L., Huehne K., Kroeckel L.;
RT "Purification and cloning of sakacin 674, a bacteriocin from
RT Lactobacillus sake Lb674.";

```

```
RL FEWS Microbiol. Lett. 115:143-150(1994).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Lb674;
RX MEDLINE=96262715; PubMed=8704983;
RA Huehne K., Axelsson L., Holck A., Kroeckel L.;
RT "Analysis of the sakacin P gene cluster from Lactobacillus sake Lb674
RL and its expression in sakacin-negative Lb. sake strains.";
RL Microbiology 142:1437-1448(1996).
RN [4]
RP PROTEIN SEQUENCE OF 19-59, AND CHARACTERIZATION.
RC STRAIN=LTH673;
RA Tichaczek P.S., Nissen-Meyer J., Nes I.F., Vogel R.F., Hammes W.P.;
RT "Characterization of the bacteriocins curvacin A from Lactobacillus
RL curvatus LTH1174 and sakacin P from L. sake LTH673.";
RL Synt. Appl. Microbiol. 15:460-465(1992).
CC -1- FUNCTION: Bactericidal activity; inhibits closely related
CC Lactobacilli, Listeria monocytogenes and Ivanovii, Enterococcus
CC faecalis, Carnobacterium sp and Brochothrix thermosphacta.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the bacteriocin class IIA/YGNV family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X75081; CAA52974.1; -; Genomic DNA.
DR EMBL; Z25816; CAA81064.1; -; Genomic DNA.
DR EMBL; Z48542; CAA88428.1; -; Genomic DNA.
DR EMBL; AF002276; AAB93970.1; -; Genomic DNA.
DR PIR; S57911; S38508.
DR PDB; 1OG7; NMR; A=19-61.
DR PDB; 1OHM; NMR; A=19-61.
DR PDB; 1OHN; NMR; A=19-61.
DR InterPro; IPR002633; Bacteriocin_II.
DR Pfam; PF01721; Bacteriocin_II; 1.
DR ProDom; PD004452; Bacteriocin_II; 1.
KW 3D-structure; Antibiotic; Antimicrobial; Bacteriocin;
KW Direct protein sequencing.
FT PROPEP 1 18
FT CHAIN 19 61 Bacteriocin sakacin P.
FT DISULFID 27 32 By similarity.
SQ SEQUENCE 61 AA; 6385 MW; B7BF14DCDD28A73D CRC64;

Query Match 21.1%; Score 64; DB 1; Length 61;
Best Local Similarity 36.7%; Pred. No. 1.8;
Matches 18; Conservative 4; Mismatches 9; Indels 18; Gaps 4;

Qy 3 YYGTNGVHCTKNSL---WGKVLKNMKYDQNTTYMGRL-QDILGWATG 47
||| ||||| | : : : : :
Db 20 YYG-NGVHCGRSCTVDWG-----TAIGNIGNNAAANWATG 54

RESULT 8
Q7B5F5 LACSK PRELIMINARY; PRT; 61 AA.
AC Q7B5F5
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Sakacin P (SppA).
GN Name=sppA;
OS Lactobacillus sakei.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1599;
[1]
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MI401;
RX MEDLINE=94012314; PubMed=8407671;
RA Larsen A.G., Vogensen F.K., Josephsen J.;
```

```
RT "Antimicrobial activity of lactic acid bacteria isolated from sour
RN doughs: purification and characterization of bavaricin A, a
RL bacteriocin produced by Lactobacillus bavaricus MI401.";
RL J. Appl. Bacteriol. 75:113-122(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MI401;
RA Johansen A.H., Vogensen F.K.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1151;
RA Ureo R., Rantsiou K., Cantoni C., Comi G., Luca C.;
RT "Sequencing and expression analysis of the sakacin P bacteriocin
RL produced by a Lactobacillus sakei strain isolated from naturally
RT fermented sausages.";
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF526262; AAM88858.1; -; Genomic DNA.
DR EMBL; AY875983; AAW79057.1; -; Genomic DNA.
SQ SEQUENCE 61 AA; 6385 MW; B7BF14DCDD28A73D CRC64;

Query Match 21.1%; Score 64; DB 2; Length 61;
Best Local Similarity 36.7%; Pred. No. 1.8;
Matches 18; Conservative 4; Mismatches 9; Indels 18; Gaps 4;

Qy 3 YYGTNGVHCTKNSL---WGKVLKNMKYDQNTTYMGRL-QDILGWATG 47
||| ||||| | : : : : :
Db 20 YYG-NGVHCGRSCTVDWG-----TAIGNIGNNAAANWATG 54

RESULT 9
Q51P75 MAGGR PRELIMINARY; PRT; 714 AA.
AC Q51P75
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG0601.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes; incertae sedis; Magnaportheaceae; Magnaporthe.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Ambruster J., Bachanteang P., Baldwin J., Barry A.,
RA Bayul T., Blitshstein B., Bloom I., Blye J., Boguslavsky L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collamore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA Dorjee K., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnrke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Hueby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamyssealis M., Karlsson E.,
RA Kells C., Kieu A., Kisher P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., Medonough S., Mcghee T., Meldrim J., Maneus L.,
RA Mesirov J., Mihaliev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pignani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
```

```

RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutnan M., Schupbach R., Seaman C., Settupalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tesmo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU01001430; BAA48943.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 714 AA; 81698 MW; FE74ED11133DCAF3 CRC64;

Query Match 21.1%; Score 64; DB 2; Length 714;
Best Local Similarity 34.2%; Pred. No. 29;
Matches 13; Conservative 7; Mismatches 18; Indels 0; Gaps 0;

QY 7 NGVHCTKSLWGKVRLLKMKYDQNTYMGRLQDILLGW 44
DB 626 NNVHCTLSNDGTLFKSSLSDHFYQMFVGRSDTIHW 663

RESULT 10
Q8YK80 ANASP PRELIMINARY; PRT; 590 AA.
AC Q8YK80;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE All8037 protein.
GN OrderedLocusNames=all8037;
OS Anabaena sp. (strain PCC 7120).
OG Plasmid pCC7120gamma.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RA "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213 (2001).
DR EMBL; AP003603; BAB77367.1; -; Genomic_DNA.
DR PIR; AF2555; AF2555.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0009291; P:unidirectional conjugation; IEA.
SQ Complete proteome; Plasmid.
SQ SEQUENCE 590 AA; 64933 MW; 7ECA68DF6FE3B469 CRC64;

Query Match 20.7%; Score 63; DB 2; Length 590;
Best Local Similarity 35.3%; Pred. No. 31;
Matches 18; Conservative 8; Mismatches 23; Indels 0; Gaps 1;

```

```

QY 3 YGTNGVHCTKSLWGKVRLLKMKYDQNTYMGRLQDILLGWATGAFGKTF 53
DB 88 YIGT--PRGTFQVGVGNKRITNPEDKRLYLPDVQRGILVSGSGSGKTF 136

RESULT 11
Q8RIM1 FUSNN PRELIMINARY; PRT; 1582 AA.
AC Q8RIM1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fusobacterium outer membrane protein family.
GN OrderedLocusNames=FN1554;
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 25586; PubMed=11889109;
RX MEDLINE=21886394; PubMed=11889109;
RA DOI=10.1128/JB.184.7.2005-2018.2002;
RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fonstein M., Kyrpides N.C., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018 (2002).
DR EMBL; AE009951; AAL93680.1; -; Genomic_DNA.
DR InterPro; IPR005546; AutoTransphbeta.
DR Pfam; PF03797; Autotransporter; 1.
DR PROSITE; PS00050; RIBOSOMAL_L23; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1582 AA; 167889 MW; 666C38783A00EB4F CRC64;

Query Match 20.7%; Score 63; DB 2; Length 1582;
Best Local Similarity 35.2%; Pred. No. 94;
Matches 19; Conservative 6; Mismatches 21; Indels 8; Gaps 3;

QY 4 YGTNGVHCTKSLWGKVRLLKMKYDQNTYMGRLQDILL---GWATGAFGKTF 53
DB 1310 FGTNGEYKTDTA--GVIDYKNHAY--GVAYVHNEIDKLRGTGWTGIVHNTF 1359

RESULT 12
Q8RHH7 FUSNN PRELIMINARY; PRT; 1630 AA.
AC Q8RHH7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fusobacterium outer membrane protein family.
GN OrderedLocusNames=FN2047;
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA DOI=10.1128/JB.184.7.2005-2018.2002;
RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fonstein M., Kyrpides N.C., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium

```

```
RT nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AE009951; AAL94131.1; -; Genomic_DNA.
DR InterPro; IPR005546; Auto_transptbeta.
DR Pfam; PF03797; Autotransporter; 1.
KW Complete proteome.
SQ SEQUENCE 1630 AA; 172743 MW; 38E9CE42F037B75A CRC64;

Query Match 20.7%; Score 63; DB 2; Length 1630;
Best Local Similarity 35.2%; Pred. No. 97;
Matches 19; Conservative 6; Mismatches 21; Indels 8; Gaps 3;

QY 4 YGTNGVHCTKSLWGVRLKNKYDQNTTYMGRQLQDILL----GWATGAFGKTF 53
Db 1358 FGTNGEYKTDTA--GVIDYKHAY--GVAYVHENEIDKLGRIGWYTGIVHNTF 1407

RESULT 13
GLNA_METVO
ID GLNA_METVO STANDARD; PRT; 446 AA.
AC P21154;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase) (GS).
GN Name=glNA;
OS Methanococcus voltae.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanococcaceae; Methanococcus.
OX NCBI_TaxID=2188;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 1537 / PS;
RX MEDLINE=90139872; PubMed=2575777; DOI=10.1016/0923-2508(89)90012-0;
RA Posset O., Sibold L., Aubert J.-P.;
RT "Nucleotide sequence and expression of the glutamine synthetase structural gene, glNA, of the archaeobacterium Methanococcus voltae.";
RL Res. Microbiol. 140:355-371(1989).
CC -!- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate + L-glutamine.
CC -!- ENZYME REGULATION: The activity of this enzyme is controlled by adenylation under conditions of abundant glutamine. The fully adenylation enzyme complex is inactive (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glutamine synthetase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
CC -----
DR EMBL; X53509; CAA37585.1; -; Genomic_DNA.
DR PIR; A43995; A43995.
DR HSP; P06201; ILGR.
DR InterPro; IPR008147; Gln_synt_beta.
DR InterPro; IPR008146; Gln_synt_C.
DR InterPro; IPR004809; GlnA.
DR Pfam; PF00120; Gln-synt_C; 1.
DR Pfam; PF03951; Gln-synt_N; 1.
DR ProDom; PD001057; Gln_synt_C; 1.
DR TIGRFAMs; TIGR00653; GlnA; 1.
DR PROSITE; PS00180; GLNA_1; 1.
DR PROSITE; PS00181; GLNA_ATP; 1.
KW Ligase
FT BINDING 396 AMP (covalent) (By similarity).
SQ SEQUENCE 446 AA; 50199 MW; 95449E4DE8542690 CRC64;

Query Match 20.6%; Score 62.5; DB 1; Length 446;
Best Local Similarity 33.9%; Pred. No. 26;
Matches 19; Conservative 9; Mismatches 17; Indels 11; Gaps 5;

QY 1 KTYYGTN--GVHCTKNSLW--GKVLKNKYDQNTTYMGRQLQDILLGWATGFGKT 52
```

```
Db 237 KPFFGMGSGMHCNQ-SIWLDGKPSF-----YDENNAH--QLSDICLSYIGILEHT 285

RESULT 14
Q8VM63_BACTA
ID Q8VM63_BACTA PRELIMINARY; PRT; 558 AA.
AC Q8VM63;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cry39ORF2 protein.
GN Name=Cry3orf2;
OS Bacillus thuringiensis (subsp. aizawai).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1433;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ito T., Sahara K., Bando H., Asano S.;
RT "Cloning and Expression of Novel Crystal Protein Genes cry39A and cry39f2 from Bacillus thuringiensis subsp. aizawai Bn1-14 Encoding Mosquitocidal Proteins.";
RL J. Insect Biotechnol. Sericology 71:123-128(2002).
DR EMBL; AB074413; BAB72017.1; -; Genomic_DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
SQ SEQUENCE 558 AA; 63360 MW; 26F6D5A013834207 CRC64;

Query Match 20.6%; Score 62.5; DB 2; Length 558;
Best Local Similarity 27.7%; Pred. No. 34;
Matches 13; Conservative 8; Mismatches 17; Indels 9; Gaps 2;

QY 9 VHCTKNS-----LMGKVR-LKNMKYDQNTTYMGRQLQDILLGWAT 46
Db 26 IECHSNEHSSKEENMLDVEVKAQQLSWSRLNLYGDFEDVSNGWKT 72

RESULT 15
Q9PZ03_GVXN
ID Q9PZ03_GVXN PRELIMINARY; PRT; 469 AA.
AC Q9PZ03;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF40.
GN Name=ORF40;
OS Xestia c-nigrum granulosis virus (XnGV) (Xestia c-nigrum granulovirus).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
OX NCBI_TaxID=51677;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hayakawa T., Ko R., Okano K., Seong S.-I., Goto C., Maeda S.;
RT "Sequence analysis of the Xestia c-nigrum granulovirus genome.";
RL Virology 262:277-297(1999).
DR EMBL; AF162221; AAF05154.1; -; Genomic_DNA.
DR HSSP; P03956; 1CGL.
DR MEROPS; M10.032; -.
DR GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000585; Hemopeptin.
DR InterPro; IPR006026; Peptidase_M.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR Pfam; PF00413; Peptidase_M10; 1.
DR PRINTS; PR00138; MATRIXIN.
DR SMART; SM00235; ZnMC; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 1.
SQ SEQUENCE 469 AA; 53808 MW; 1BF4EB5C210E241 CRC64;

Query Match 20.2%; Score 61.5; DB 2; Length 469;
```

Best Local Similarity 35.9%; Pred. No. 38;
Matches 14; Conservative 9; Mismatches 7; Indels 9; Gaps 2;
Qy 5 GTNGVHCTKNSLWGVRLKKNMKNKYDONTTMM--GRLQDIL 41
Db 335 GDEIIATRNLLW-----YEHKKNSTLMNVGRVQDYL 366

Search completed: May 2, 2006, 17:30:32
Job time : 237 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 17:30:50 ; Search time 46 Seconds
(without alignments)
97.054 Million cell updates/sec

Title: US-10-644-927-1

Perfect score: 304

Sequence: 1 KTYGTGNGVHCTKNSLWGVK.....GRLODILLGNATGAFGKTFH 54

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5 COMB.pap.*

2: /cgn2_6/ptodata/1/iaa/6 COMB.pap.*

3: /cgn2_6/ptodata/1/iaa/H COMB.pap.*

4: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pap.*

5: /cgn2_6/ptodata/1/iaa/RE COMB.pap.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 69.5 | 22.9 | 420 | 2 | US-09-328-352-5261 |
| 2 | 64 | 21.1 | 61 | 2 | US-09-068-507C-5 |
| 3 | 62 | 20.4 | 85 | 2 | US-09-270-767-44549 |
| 4 | 60 | 19.7 | 36 | 2 | US-08-924-629C-69 |
| 5 | 60 | 19.7 | 37 | 2 | US-09-030-619-209 |
| 6 | 60 | 19.7 | 37 | 2 | US-09-444-281-90 |
| 7 | 60 | 19.7 | 61 | 2 | US-08-924-629C-68 |
| 8 | 60 | 19.7 | 300 | 2 | US-09-982-616-9 |
| 9 | 59.5 | 19.6 | 577 | 2 | US-09-248-796A-18807 |
| 10 | 59 | 19.4 | 62 | 2 | US-08-924-629C-71 |
| 11 | 56.5 | 18.6 | 242 | 2 | US-09-543-681A-4928 |
| 12 | 56 | 18.4 | 66 | 2 | US-09-107-532A-4680 |
| 13 | 56 | 18.4 | 909 | 2 | US-09-982-616-11 |
| 14 | 56 | 18.4 | 10182 | 2 | US-09-134-001C-3159 |
| 15 | 55.5 | 18.3 | 243 | 2 | US-09-216-295-13 |
| 16 | 55.5 | 18.3 | 244 | 2 | US-09-632-570-13 |
| 17 | 55.5 | 18.3 | 244 | 2 | US-09-632-575-43 |
| 18 | 55 | 18.1 | 41 | 2 | US-08-924-629C-70 |
| 19 | 55 | 18.1 | 177 | 2 | US-09-248-796A-18521 |
| 20 | 54.5 | 17.9 | 494 | 2 | US-08-378-313-23 |
| 21 | 54.5 | 17.9 | 494 | 2 | US-08-378-313-29 |
| 22 | 54.5 | 17.9 | 641 | 2 | US-09-071-035-456 |
| 23 | 54.5 | 17.9 | 641 | 2 | US-10-206-576-456 |
| 24 | 54.5 | 17.9 | 1313 | 2 | US-09-071-035-456 |
| 25 | 54.5 | 17.9 | 1313 | 2 | US-09-071-035-454 |
| 26 | 54.5 | 17.9 | 1313 | 2 | US-10-206-576-450 |
| 27 | 54.5 | 17.9 | 1313 | 2 | US-10-206-576-454 |

| | | | | | | |
|----|------|------|------|---|----------------------|--------------------|
| 28 | 54 | 17.8 | 697 | 2 | US-09-489-039A-7485 | Sequence 7485, Ap |
| 29 | 53.5 | 17.6 | 56 | 2 | US-09-834-309-3 | Sequence 3, Appli |
| 30 | 53.5 | 17.6 | 74 | 2 | US-09-107-532A-5289 | Sequence 5289, Ap |
| 31 | 53.5 | 17.6 | 134 | 2 | US-09-834-309-4 | Sequence 4, Appli |
| 32 | 53.5 | 17.6 | 1012 | 2 | US-08-126-505A-15 | Sequence 15, Appli |
| 33 | 53.5 | 17.6 | 1033 | 2 | US-09-834-309-1 | Sequence 1, Appli |
| 34 | 53 | 17.4 | 233 | 1 | US-08-032-848C-12 | Sequence 12, Appli |
| 35 | 53 | 17.4 | 263 | 2 | US-09-216-295-24 | Sequence 24, Appli |
| 36 | 53 | 17.4 | 264 | 2 | US-09-632-570-24 | Sequence 24, Appli |
| 37 | 53 | 17.4 | 264 | 2 | US-09-632-575-54 | Sequence 54, Appli |
| 38 | 53 | 17.4 | 384 | 2 | US-09-311-170-2 | Sequence 2, Appli |
| 39 | 53 | 17.4 | 929 | 2 | US-09-134-000C-6424 | Sequence 6424, Ap |
| 40 | 52.5 | 17.3 | 371 | 2 | US-09-248-796A-17748 | Sequence 17748, A |
| 41 | 52.5 | 17.3 | 418 | 2 | US-09-610-104C-12 | Sequence 12, Appli |
| 42 | 52.5 | 17.3 | 911 | 2 | US-09-356-952-6 | Sequence 6, Appli |
| 43 | 52.5 | 17.3 | 1176 | 1 | US-08-257-999-2 | Sequence 2, Appli |
| 44 | 52 | 17.1 | 219 | 2 | US-09-540-236-2633 | Sequence 2633, Ap |
| 45 | 52 | 17.1 | 221 | 2 | US-09-270-767-33753 | Sequence 33753, A |

ALIGNMENTS

RESULT 1

US-09-328-352-5261

; Sequence 5261, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/328.352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 5261

; LENGTH: 420

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-5261

Query Match 22.9%; Score 69.5; DB 2; Length 420;

Best Local Similarity 38.5%; Pred. No. 0.47;

Matches 20; Conservative 8; Mismatches 13; Indels 11; Gaps 4;

QY 1 KTYGTG---GVHCTKNSLWGVKVKMKY--DQNTYMGRLQDI---LLGW 44

DB 361 ESVVVTNPKSGKSHNKANIIWK---NNTPYSGDRDDTYRRLRDITSPLIGW 409

RESULT 2

US-09-068-507C-5

; Sequence 5, Application US/09068507C

; Patent No. 6790951

; GENERAL INFORMATION:

; APPLICANT: ELJISINK, VINCENT et al.

; TITLE OF INVENTION: EXPRESSION SYSTEM IN MICROORGANISM AND ITS USE FOR EXPRESSING

; FILE REFERENCE: HETEROLOGOUS AND HOMOLOGOUS PROTEINS

; CURRENT APPLICATION NUMBER: US/09/068.507C

; CURRENT FILING DATE: 1998-07-15

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 5

; LENGTH: 61

; TYPE: PRT

; ORGANISM: Lactobacillus sake

US-09-068-507C-5

Query Match 21.1%; Score 64; DB 2; Length 61;

Best Local Similarity 36.7%; Pred. No. 0.23;

Matches 18; Conservative 4; Mismatches 9; Indels 18; Gaps 4;


```

; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18807
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18807

Query Match      19.6%; Score 59.5; DB 2; Length 577;
Best Local Similarity 30.6%; Pred. No. 16; Gaps 2;
Matches 15; Conservative 7; Mismatches 12; Indels 15; Gaps 2;

QY       6 TNGVCHTKNSLWGCKVRLLKMKKYDNTTMYMGRQLDLILLGWATGAFGKTFFH 54
          ||| :||| |:::||::| :|||
DB     261 TTGVQCDESLWKL--IPNLKHLNN-----OTSAICKFTH 294

RESULT 10
US-08-924-629C-71
; Sequence 71, Application US/08924629C
; Patent No. 6403082
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vederas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Worobo, Randy W.
; APPLICANT: Worobo, Rodney J.
; APPLICANT: Greer, G. Gordon
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Aision
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: No. 6403082elBacteriocins, Transport and Vector System
; FILE REFERENCE: 660.000SUS
; CURRENT APPLICATION NUMBER: US/08/924.629C
; CURRENT FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Pediocin PAI
US-08-924-629C-71

Query Match      19.4%; Score 59; DB 2; Length 62;
Best Local Similarity 37.5%; Pred. No. 1.1;
Matches 18; Conservative 5; Mismatches 9; Indels 16; Gaps 4

QY       3 YYGTNNGVHCYKNLSI---WGKRVRLKKMKYDNTTYMGLRQLDLLGWATG 47
          ||| :||| |:::||::| :|||
DB     20 YYG-NGYTCGHSCSVDMGKA-----TTCI--INNGAMAWATG 54

RESULT 11
US-09-543-681A-4928
; Sequence 4928, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEINS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128, 706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4928
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; FEATURE:
```

```
; NAME/KEY: UNSURE
; LOCATION: (18)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-543-681A-4928

Query Match      18.6%   Score 56.5; DB 2; Length 242;
Best Local Similarity 28.1%; Pred.No.14;
Matches 18; Conservative 8; Mismatches 21; Indels 17; Gaps 3;

QY    4 YGNGVHCTKNSLWGVRLKN-----MKYDQNTTYMGRGLDLILGWATGAF 49
      |||:|||||||
Db     106 YKTSSTDSAKISVFGLVNKNTGEIGGYIIDVIDIKVDNKRAYLGRVADKI--W-TGSS 162
      |||:|||||||

QY    50 GKTF 53
      |
Db     163 ATIF 166

RESULT 12
US-09-107-532A-4680
; Sequence 4680, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; FILING DATE: 30-Jun-1998
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4680:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 66 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...66
; SEQUENCE DESCRIPTION: SEQ ID NO: 4680:
US-09-107-532A-4680
```

```

Qy 3 YGTNGVHCTKNSL---WGK 19
   ||| |||:||||
Db 26 YYG-NGVYCTKTKRKTVDMAK 44

RESULT 13
US-09-982-616-11
; Sequence 11, Application US/09982616
; Patent No. 6955884
; GENERAL INFORMATION:
; APPLICANT: Frances E. Lund
; APPLICANT: Troy D. Randall
; APPLICANT: Santiago Partida-Sanchez
; TITLE OF INVENTION: CD38 MODULATED CHEMOTAXIS
; FILE REFERENCE: AP33438 068443.0106
; CURRENT APPLICATION NUMBER: US/09/982,616
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 60/241,065
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 909
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Reverse translation of SM38
US-09-982-616-11

Query Match 18.4%; Score 56; DB 2; Length 909;
Best Local Similarity 32.7%; Pred. No. 85;
Matches 16; Conservative 2; Mismatches 15; Indels 16;

Qy 4 YGTNGVHCTKNSLWGVKVLKMKYDQNTTYMGRLODILLG--WATGAFG 50
   ||||| :|
Db 477 YGTNGTNC-----YGCNTTYGGCARWSNGCNSGNGC 511

RESULT 14
US-09-134-001C-3159
; Sequence 3159, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3159
; LENGTH: 10182
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3159

Query Match 18.4%; Score 56; DB 2; Length 10182;
Best Local Similarity 46.7%; Pred. No. 1.8e+03;
Matches 14; Conservative 2; Mismatches 14; Indels 0;

Qy 6 TNGVHCTKNSLWGVKVLKMKYDQNTTYMG 35
   ||| |||:||||
Db 4753 TQVNTTKNDLNGNDKLAERKDANTTIDG 4782

RESULT 15
US-09-216-295-13
; Sequence 13, Application US/09216295
; Patent No. 6268328

```

GENERAL INFORMATION:
; APPLICANT: Mitchinson, Colin
; APPLICANT: Wendt, Dan J.
; TITLE OF INVENTION: No. 6268328el Variant EGIII-Like Cellulase Compositions
; FILE REFERENCE: GC555
; CURRENT APPLICATION NUMBER: US/09/216,295
; CURRENT FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Fusarium javanicum (1)
US-09-216-295-13

Query Match 18.3%; Score 55.5; DB 2; Length 243;
Best Local Similarity 31.4%; Pred. No. 19;
Matches 16; Conservative 8; Mismatches 20; Indels 7; Gaps 2;
Qy 6 TNGVHCTKNSLWKGKRLKNMKYDONTTMYGRLQDILGWA-----TGAFGK 51
Db 34 TAGAYTIYNLWKGKNAES--GEQCTTNSGEQSDGSIAWSVSWTGGQQG 82

Search completed: May 2, 2006, 17:32:07
Job time : 47 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 17:42:01 ; Search time 164 Seconds
(without alignments)
137.578 Million cell updates/sec

Title: US-10-644-927-1
Perfect score: 304
Sequence: 1 KTYGTNGVHCTKNSLWKV.....GRLQDILLGWATGFGKTFH 54

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 304 | 100.0 | 54 | 5 | US-10-644-927-1 |
| 2 | 62.5 | 20.6 | 558 | 5 | US-10-781-979-27 |
| 3 | 60 | 19.7 | 36 | 3 | US-09-883-343A-69 |
| 4 | 60 | 19.7 | 37 | 3 | US-09-030-619-209 |
| 5 | 60 | 19.7 | 37 | 3 | US-09-912-609-97 |
| 6 | 60 | 19.7 | 37 | 4 | US-10-277-232-209 |
| 7 | 60 | 19.7 | 37 | 4 | US-10-277-233-209 |
| 8 | 60 | 19.7 | 61 | 3 | US-09-883-343A-68 |
| 9 | 60 | 19.7 | 300 | 3 | US-09-982-616-9 |
| 10 | 60 | 19.7 | 300 | 3 | US-09-773-307B-2 |
| 11 | 60 | 19.7 | 300 | 4 | US-10-295-027-1348 |
| 12 | 60 | 19.7 | 300 | 4 | US-10-454-238-2 |
| 13 | 60 | 19.7 | 300 | 4 | US-10-322-696-162 |
| 14 | 60 | 19.7 | 300 | 5 | US-10-753-267-34 |
| 15 | 60 | 19.7 | 688 | 5 | US-10-450-763-37242 |
| 16 | 59 | 19.4 | 62 | 3 | US-09-883-343A-71 |
| 17 | 59 | 19.4 | 399 | 4 | US-10-369-493-6141 |
| 18 | 59 | 19.4 | 753 | 5 | US-10-983-198-42 |
| 19 | 58.5 | 19.2 | 95 | 4 | US-10-424-599-280304 |
| 20 | 58 | 19.1 | 511 | 4 | US-10-424-599-213009 |
| 21 | 57.5 | 18.9 | 461 | 4 | US-10-410-432-10 |
| 22 | 57.5 | 18.9 | 461 | 4 | US-10-669-174-10 |
| 23 | 57.5 | 18.9 | 461 | 5 | US-10-984-956A-10 |
| 24 | 57.5 | 18.9 | 461 | 5 | US-10-732-923-19385 |
| 25 | 57.5 | 18.9 | 461 | 6 | US-11-097-143-40056 |
| 26 | 57 | 18.8 | 564 | 4 | US-10-424-599-19752 |
| 27 | 57 | 18.8 | 581 | 4 | US-10-282-122A-72532 |

| | | | | | | |
|----|------|------|-------|---|----------------------|-------------------|
| 28 | 56 | 18.4 | 636 | 4 | US-10-424-599-225665 | Sequence 225665, |
| 29 | 56 | 18.4 | 909 | 3 | US-09-982-616-11 | Sequence 11, Appl |
| 30 | 56 | 18.4 | 6641 | 4 | US-10-282-122A-70580 | Sequence 70580, A |
| 31 | 56 | 18.4 | 10203 | 4 | US-10-661-809-23 | Sequence 23, Appl |
| 32 | 56 | 18.4 | 10203 | 4 | US-10-724-972A-4098 | Sequence 4098, Ap |
| 33 | 55.5 | 18.3 | 81 | 4 | US-10-425-115-338836 | Sequence 338836, |
| 34 | 55.5 | 18.3 | 108 | 4 | US-10-425-115-304001 | Sequence 304001, |
| 35 | 55.5 | 18.3 | 244 | 4 | US-10-441-626-13 | Sequence 13, Appl |
| 36 | 55.5 | 18.3 | 244 | 4 | US-10-441-626-13 | Sequence 70314, A |
| 37 | 55.5 | 18.3 | 379 | 4 | US-10-425-114-70314 | Sequence 251103, |
| 38 | 55.5 | 18.3 | 553 | 4 | US-10-425-115-251103 | Sequence 60740, A |
| 39 | 55.5 | 18.3 | 608 | 4 | US-10-425-114-60740 | Sequence 70, Appl |
| 40 | 55 | 18.1 | 41 | 3 | US-09-883-343A-70 | Sequence 5582, Ap |
| 41 | 55 | 18.1 | 112 | 3 | US-09-738-626-5582 | Sequence 203257, |
| 42 | 55 | 18.1 | 153 | 4 | US-10-425-115-203257 | Sequence 55991, A |
| 43 | 55 | 18.1 | 285 | 4 | US-10-282-122A-55991 | Sequence 103806, |
| 44 | 55 | 18.1 | 2071 | 4 | US-10-437-963-103806 | Sequence 27507, A |
| 45 | 54.5 | 17.9 | 302 | 6 | US-11-097-143-27507 | |

ALIGNMENTS

RESULT 1

US-10-644-927-1
; Sequence 1, Application US/10644927
; Publication No. US20050153881A1
; GENERAL INFORMATION:
; APPLICANT: Stern, Norman J
; APPLICANT: Svetoich, Edward A.
; APPLICANT: Eruslanov, Boris V.
; APPLICANT: Volodina, Larisa I.
; APPLICANT: Kovalev, Yuri N.
; APPLICANT: Kudryavtseva, Tamara Y.
; APPLICANT: Pereygin, Vladimir V.
; APPLICANT: Pokhilenko, Victor D.
; APPLICANT: Levchuk, Vladimir P.
; APPLICANT: Borzenkov, Valery N.
; APPLICANT: Svetoich, Olga E.
; APPLICANT: Mitsevich, Eugeni V.
; APPLICANT: Mitsevich, Irina P.
; TITLE OF INVENTION: Bacteriocins and Novel Bacterial Strains
; FILE REFERENCE: D.N. 0135.03
; CURRENT APPLICATION NUMBER: US/10/644,927
; CURRENT FILING DATE: 2003-08-21
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Lactobacillus salivarius
US-10-644-927-1

Query Match 100.0%; Score 304; DB 5; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.4e-31;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KTYGTNGVHCTKNSLWKVRLKNMKYDONTYMGRLQDILLGWATGFGKTFH 54
Db 1 KTYGTNGVHCTKNSLWKVRLKNMKYDONTYMGRLQDILLGWATGFGKTFH 54

RESULT 2

US-10-781-979-27
; Sequence 27, Application US/10781979
; Publication No. US20040250311A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-008, A Delta-Endotoxin Gene and

; TITLE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274147
; CURRENT APPLICATION NUMBER: US/10/781.979
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: 60/448,797
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-781-979-27

Query Match 20.6%; Score 62.5; DB 5; Length 558;
Best Local Similarity 27.7%; Pred. No. 32;
Matches 13; Conservative 8; Mismatches 17; Indels 9; Gaps 2;
Qy 9 VHCYKNS-----LNGKVR-LKNMKYDQNTTVMGRLODILLGWAT 46
Db 26 IECMSNEHSSKEEMLMWDEVKQKLSWSRLLYNGDPEDVSNWGT 72

RESULT 3

US-09-883-343A-69
; Sequence 69, Application US/09883343A
; Publication No. US20030039632A1
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vederas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Worobo, Randy W.
; APPLICANT: Worobo, Rodney J.
; APPLICANT: Greer, G. Gordon
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Alison
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: No. US20030039632A1elbacteriocins, Transport and Vector System an
; FILE REFERENCE: 660.000505
; CURRENT APPLICATION NUMBER: US/09/883.343A
; CURRENT FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US/08/924,629
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Mesenteriocin Y105
US-09-883-343A-69

Query Match 19.7%; Score 60; DB 3; Length 36;
Best Local Similarity 48.4%; Pred. No. 3;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;
Qy 3 YYGTNGVHCTKNSL---WGKV-----RLKN 24
Db 2 YYG-NGVHCTKSGCSVNWGEAASAGIHRLAN 31

RESULT 4

US-09-010-619-209
; Sequence 209, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfle, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.

; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 209
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Leuconostoc gelidium
US-09-030-619-209

Query Match 19.7%; Score 60; DB 3; Length 37;
Best Local Similarity 48.4%; Pred. No. 3.1;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;
Qy 3 YYGTNGVHCTKNSL---WGKV-----RLKN 24
Db 2 YYG-NGVHCTKSGCSVNWGEAASAGVHRLAN 31

RESULT 5

US-09-912-609-97
; Sequence 97, Application US/09912609
; Publication No. US20020041898A1
; GENERAL INFORMATION:
; APPLICANT: UNGER, EVAN C.
; APPLICANT: MATSUNAGA, TERRY ONICHI
; APPLICANT: ROMANOWSKI, MAREK J.
; APPLICANT: ROMANOWSKI, MAREK J.
; TITLE OF INVENTION: NOVEL TARGETED DELIVERY SYSTEMS FOR BIOACTIVE AGENTS
; FILE REFERENCE: 5030-0001.24
; CURRENT APPLICATION NUMBER: US/09/912,609
; CURRENT FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 09/703,474
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 09/478,124
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-912-609-97

Query Match 19.7%; Score 60; DB 3; Length 37;
Best Local Similarity 48.4%; Pred. No. 3.1;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;
Qy 3 YYGTNGVHCTKNSL---WGKV-----RLKN 24
Db 2 YYG-NGVHCTKSGCSVNWGEAASAGVHRLAN 31

RESULT 6

US-10-277-232-209
; Sequence 209, Application US/10277232
; Publication No. US20030211537A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfle, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406C1
; CURRENT APPLICATION NUMBER: US/10/277,232
; CURRENT FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 209
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Leuconostoc gelidum
US-10-277-232-209

Query Match 19.7%; Score 60; DB 4; Length 37;
Best Local Similarity 48.4%; Pred. No. 3.1;
Matches 15; Conservative 2; Mismatches 3;

QY 3 YGTNGVHCTKNSL---WGKV-----RLKN 24
DB 2 YYG-NGVHCTKSGCSVNWGEAFSAGVHRLAN 31

RESULT 7
US-10-277-233-209
; Sequence 209, Application US/10277233
; Publication No. US20030232750A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erle, Douglas R.
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406C1
; CURRENT APPLICATION NUMBER: US/10/277,233
; CURRENT FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 209
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Leuconostoc gelidum
US-10-277-233-209

Query Match 19.7%; Score 60; DB 4; Length 37;
Best Local Similarity 48.4%; Pred. No. 3.1;
Matches 15; Conservative 2; Mismatches 3;

QY 3 YGTNGVHCTKNSL---WGKV-----RLKN 24
DB 2 YYG-NGVHCTKSGCSVNWGEAFSAGVHRLAN 31

RESULT 8
US-09-883-343A-68
; Sequence 68, Application US/09883343A
; Publication No. US20030039632A1
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vederas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Worobo, Randy W.
; APPLICANT: Worobo, Rodney J.
; APPLICANT: Greer, G. Gordon
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Alson
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: No. US20030039632A1elBacteriocins, Transport and Vector System an
; FILE REFERENCE: 660.0005US

; CURRENT APPLICATION NUMBER: US/09/883,343A
; CURRENT FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US/08/924,629
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Leucocin A
US-09-883-343A-68

Query Match 19.7%; Score 60; DB 3; Length 61;
Best Local Similarity 48.4%; Pred. No. 5.5;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;

QY 3 YGTNGVHCTKNSL---WGKV-----RLKN 24
DB 26 YYG-NGVHCTKSGCSVNWGEAFSAGVHRLAN 55

RESULT 9
US-09-982-616-9
; Sequence 9, Application US/09982616
; Publication No. US20020127646A1
; GENERAL INFORMATION:
; APPLICANT: Frances E. Lund
; APPLICANT: Troy D. Randall
; APPLICANT: Santiago Partida-Sanchez
; TITLE OF INVENTION: CD38 MODULATED CHEMOTAXIS
; FILE REFERENCE: AP33438 068443.0106
; CURRENT APPLICATION NUMBER: US/09/982,616
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 60/241,065
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-982-616-9

Query Match 19.7%; Score 60; DB 3; Length 300;
Best Local Similarity 35.0%; Pred. No. 33;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

QY 5 GTNGVHCTKNSLWCKVRLKNMKYDQNTTYMGRLODILLGW 44
DB 113 GTQVPCNKILLWSRIKDLAHQFTQVORDMFTLEDTLLGY 152

RESULT 10
US-09-773-307B-2
; Sequence 2, Application US/09773307B
; Publication No. US20030027134A1
; GENERAL INFORMATION:
; APPLICANT: BML, INC.
; TITLE OF INVENTION: Method of Detecting Risk Factor for Onset of Diabetes
; FILE REFERENCE: PEM37
; CURRENT APPLICATION NUMBER: US/09/773,307B
; CURRENT FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Hominidae
US-09-773-307B-2

Query Match 19.7%; Score 60; DB 3; Length 300;

Qy 5 GTNGVHCTKNSLWKGKVLKNNKMYDQNTTYMGRLODILLGW 44
Db 83 GTQVPCNKILLWSRIKDLAHOFTQVQDMFTLEDTLLGY 122

Search completed: May 2, 2006, 17:45:42
Job time : 165 secs

CURRENT APPLICATION NUMBER: US/10/753,267
CURRENT FILING DATE: 2004-01-08
PRIOR APPLICATION NUMBER: US 60/439,683
PRIOR FILING DATE: 2003-01-13
PRIOR APPLICATION NUMBER: US 60/445,216
PRIOR FILING DATE: 2003-02-05
PRIOR APPLICATION NUMBER: US 60/448,036
PRIOR FILING DATE: 2003-02-18
PRIOR APPLICATION NUMBER: US 60/454,189
PRIOR FILING DATE: 2003-03-12
PRIOR APPLICATION NUMBER: US 60/457,541
PRIOR FILING DATE: 2003-03-25
PRIOR APPLICATION NUMBER: US 60/466,411
PRIOR FILING DATE: 2003-04-29
PRIOR APPLICATION NUMBER: US 60/469,041
PRIOR FILING DATE: 2003-05-08
PRIOR APPLICATION NUMBER: US 60/477,414
PRIOR FILING DATE: 2003-06-10
PRIOR APPLICATION NUMBER: US 60/478,560
PRIOR FILING DATE: 2003-06-13
PRIOR APPLICATION NUMBER: US 60/489,772
PRIOR FILING DATE: 2003-07-24
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 130
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 34
LENGTH: 300
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-753-267-34

Query Match 19.7%; Score 60; DB 5; Length 300;
Best Local Similarity 35.0%; Pred. No. 33;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

Qy 5 GTNGVHCTKNSLWKGKVLKNNKMYDQNTTYMGRLODILLGW 44
Db 113 GTQVPCNKILLWSRIKDLAHOFTQVQDMFTLEDTLLGY 152

RESULT 15
US-10-450-763-37242
Sequence 37242, Application US/10450763
Publication No. US20050196754A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 37242
LENGTH: 688
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (49)..(136)
OTHER INFORMATION: ADP-ribosyl cyclase domain identified by Pfam, accession name
OTHER INFORMATION: Rib_hydrolyase, E-value=1.2e-88, Pfam score of 219.7
US-10-450-763-37242

Query Match 19.7%; Score 60; DB 5; Length 688;
Best Local Similarity 35.0%; Pred. No. 85;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 17:43:06 ; Search time 25 Seconds
(without alignments)
98.232 Million cell updates/sec

Title: US-10-644-927-1
Perfect score: 304
Sequence: 1 KTYGTNGVHCTKNSLWGVK.....GRLODILLGWATGAFGKTFF 54

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 232119 seqs, 45477862 residues

Total number of hits satisfying chosen parameters: 232119

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:
1: /SIDSS/prodata/2/pubpaa/US08_NEW_PUB.pep1.*
2: /SIDSS/prodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /SIDSS/prodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /SIDSS/prodata/2/pubpaa/US08_NEW_PUB.pep.*
5: /SIDSS/prodata/2/pubpaa/US09_NEW_PUB.pep.*
6: /SIDSS/prodata/2/pubpaa/US09_NEW_PUB.pep1.*
7: /SIDSS/prodata/2/pubpaa/US10_NEW_PUB.pep.*
8: /SIDSS/prodata/2/pubpaa/US10_NEW_PUB.pep1.*
9: /SIDSS/prodata/2/pubpaa/US11_NEW_PUB.pep1.*
10: /SIDSS/prodata/2/pubpaa/US11_NEW_PUB.pep1.*
11: /SIDSS/prodata/2/pubpaa/US11_NEW_PUB.pep1.*
12: /SIDSS/prodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 66 | 21.7 | 557 | 11 | US-11-188-298-21798 |
| 2 | 65.5 | 21.5 | 471 | 11 | US-11-188-298-21598 |
| 3 | 62.5 | 20.6 | 446 | 11 | US-11-188-298-13174 |
| 4 | 60 | 19.7 | 37 | 11 | US-11-088-783-90 |
| 5 | 60 | 19.7 | 300 | 11 | US-11-058-924-7 |
| 6 | 57.5 | 18.9 | 670 | 11 | US-11-188-298-18003 |
| 7 | 57 | 18.8 | 473 | 11 | US-11-188-298-15548 |
| 8 | 56.5 | 18.6 | 812 | 11 | US-11-188-298-9586 |
| 9 | 56 | 18.4 | 226 | 11 | US-11-188-298-13006 |
| 10 | 55.5 | 18.3 | 127 | 11 | US-11-188-298-15017 |
| 11 | 54.5 | 17.9 | 480 | 9 | US-10-915-002-299 |
| 12 | 53.5 | 17.6 | 56 | 9 | US-10-921-415-3 |
| 13 | 53.5 | 17.6 | 134 | 9 | US-10-921-415-4 |
| 14 | 53.5 | 17.6 | 278 | 11 | US-11-087-099-587 |
| 15 | 53.5 | 17.6 | 353 | 11 | US-11-087-099-5419 |
| 16 | 53.5 | 17.6 | 364 | 11 | US-11-087-099-3112 |
| 17 | 53.5 | 17.6 | 373 | 11 | US-11-087-099-7159 |
| 18 | 53.5 | 17.6 | 510 | 11 | US-11-188-298-394 |
| 19 | 53.5 | 17.6 | 510 | 11 | US-11-188-298-6950 |
| 20 | 53.5 | 17.6 | 510 | 11 | US-11-188-298-21258 |
| 21 | 53.5 | 17.6 | 663 | 11 | US-11-188-298-21626 |

| | | | | | |
|----|------|------|------|----|----------------------|
| 22 | 53.5 | 17.6 | 1033 | 9 | US-10-921-415-1 |
| 23 | 53.5 | 17.6 | 1033 | 9 | US-10-501-841-15 |
| 24 | 52.5 | 17.3 | 367 | 11 | US-11-096-568A-27618 |
| 25 | 52.5 | 17.3 | 385 | 11 | US-11-096-568A-27617 |
| 26 | 52.5 | 17.3 | 419 | 11 | US-11-096-568A-27616 |
| 27 | 52.5 | 17.3 | 510 | 11 | US-11-188-298-13840 |
| 28 | 52 | 17.1 | 193 | 11 | US-11-079-463-6622 |
| 29 | 52 | 17.1 | 719 | 9 | US-10-511-538-247 |
| 30 | 52 | 17.1 | 1263 | 11 | US-11-087-099-7209 |
| 31 | 51.5 | 16.9 | 381 | 11 | US-11-079-463-7731 |
| 32 | 51.5 | 16.9 | 472 | 11 | US-11-188-298-4518 |
| 33 | 51.5 | 16.9 | 510 | 11 | US-11-188-298-8654 |
| 34 | 51.5 | 16.9 | 510 | 11 | US-11-188-298-16235 |
| 35 | 51.5 | 16.9 | 511 | 11 | US-11-188-298-11478 |
| 36 | 51.5 | 16.9 | 513 | 11 | US-11-188-298-6540 |
| 37 | 51.5 | 16.9 | 513 | 11 | US-11-188-298-14116 |
| 38 | 51.5 | 16.9 | 1069 | 11 | US-11-098-686-10296 |
| 39 | 51 | 16.8 | 429 | 11 | US-11-087-099-8332 |
| 40 | 51 | 16.8 | 858 | 11 | US-11-087-099-916 |
| 41 | 51 | 16.8 | 3390 | 9 | US-10-204-252-22 |
| 42 | 50.5 | 16.6 | 355 | 11 | US-11-079-463-6473 |
| 43 | 50.5 | 16.6 | 399 | 11 | US-11-079-463-8800 |
| 44 | 50.5 | 16.6 | 444 | 11 | US-11-188-298-15511 |
| 45 | 50.5 | 16.6 | 491 | 9 | US-10-506-454-338 |

ALIGNMENTS

RESULT 1
US-11-188-298-21798
; Sequence 21798, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 21798
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(557)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-188-298-21798

Query Match 21.7%; Score 66; DB 11; Length 557;
Best Local Similarity 36.5%; Pred. NO. 0.87; Mismatches 22; Indels 2; Gaps 2;
Matches 19; Conservative 9;
QY 3 YYGNGVHCTKNSLW-GKVRLLKNMKYDQNT-TYMGRLQDILLGWATGAFGKT 52
DB 66 YETTELVENKNGTYQWVVKPKTKVIEPKTNTHVPLGVMLVGVGXGNGNST 117

RESULT 2
US-11-188-298-21598
; Sequence 21598, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31

```
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 21598
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Sulfolobus solfataricus
US-11-188-298-21598

Query Match      21.5%; Score 65.5; DB 11; Length 471;
Best Local Similarity 48.6%; Pred. No. 0.84;
Matches 17; Conservative 2; Mismatches 13; Indels 3; Gaps 2;

Qy 1 KTYVGTN--GVHCTKNSLWGVRLKNMKYDQNTTY 33
Db 255 KPIYGDNGTGMH--THLSLWTKDGKKNLWYDPNDEY 288

RESULT 3
US-11-188-298-13174
; Sequence 13174, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; PRIOR FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 13174
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Methanococcus voltae
US-11-188-298-13174

Query Match      20.6%; Score 62.5; DB 11; Length 446;
Best Local Similarity 33.9%; Pred. No. 2;
Matches 19; Conservative 9; Mismatches 17; Indels 11; Gaps 5;

Qy 1 KTYVGTN--GVHCTKNSLW--GKVLKNMKYDQNTTYMGRLODILLGWATGAFGKT 52
Db 237 KPFGMGNGSHCNQ--SIWLDGRPSF---YDENNAH--QUSDICLSYIGILEHT 285

RESULT 4
US-11-068-783-90
; Sequence 90, Application US/11068783
; Publication No. US20050260715A1
; GENERAL INFORMATION:
; APPLICANT: Burian, Jan
; TITLE OF INVENTION: EFFICIENT METHODS FOR PRODUCING
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES IN HOST CELLS
; FILE REFERENCE: 660081.411
; CURRENT APPLICATION NUMBER: US/11/068,783
; PRIOR FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: US/09/444,281
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 90
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Leuconostoc gelidum
US-11-068-783-90

Query Match      19.7%; Score 60; DB 11; Length 37;
Best Local Similarity 48.4%; Pred. No. 0.26;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;

Qy 3 YVGTNGVHCTKNSL---WGKV-----RLKN 24
Db 2 YVGTNGVHCTKSGCSVNWGEAFSGVHRLAN 31
```

```
RESULT 5
US-11-058-924-7
; Sequence 7, Application US/11058924
; Publication No. US20060019308A1
; GENERAL INFORMATION:
; APPLICANT: Lund, Frances E.
; APPLICANT: Randall, Troy D.
; TITLE OF INVENTION: CD38 MODULATED CHEMOTAXIS
; FILE REFERENCE: 13315/5
; CURRENT APPLICATION NUMBER: US/11/058,924
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 09/982,616
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 60/241,065
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-058-924-7

Query Match      19.7%; Score 60; DB 11; Length 300;
Best Local Similarity 35.0%; Pred. No. 2.7;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

Qy 5 GTNGVHCTKNSLWGVRLKNMKYDQNTTYMGRLODILLGW 44
Db 113 GTQVPCNKILLWSRIKDLAHOFTQVQDMFTLEDLLGY 152

RESULT 6
US-11-188-298-18003
; Sequence 18003, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 18003
; LENGTH: 670
; TYPE: PRT
; ORGANISM: GIBBERELLA ZEAE PH-1
US-11-188-298-18003

Query Match      18.9%; Score 57.5; DB 11; Length 670;
Best Local Similarity 29.5%; Pred. No. 14;
Matches 13; Conservative 8; Mismatches 18; Indels 5; Gaps 2;

Qy 3 YVG--TNGVHCTKNSLWGVRLKNMKYDQNTTYMGRLODILLGW 44
Db 566 YHGVAVNPiHL---SLWERLRLGLGKSYDPEADWRKAKVQDMRAEM 606

RESULT 7
US-11-188-298-15548
; Sequence 15548, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
```

```
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 15548
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Sulfolobus acidocaldarius
US-11-188-298-15548

Query Match      18.8%; Score 57; DB 11; Length 473;
Best Local Similarity 45.7%; Pred. No. 11;
Matches 16; Conservative 4; Mismatches 11; Indels 4; Gaps 3;

Qy 1 KTYGTN--GVHCTKNSLWGVRLKNMKYDQNTTY 33
Db 257 KPFFDNGSGMH--THFSLWTXDG-KNLMDPND EY 289

RESULT 8
US-11-188-298-9586
; Sequence 9586, Application US/11188298
; Publication No. US2006007522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 9586
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Eremothecium gossypii
US-11-188-298-9586

Query Match      18.6%; Score 56.5; DB 11; Length 812;
Best Local Similarity 29.4%; Pred. No. 24;
Matches 15; Conservative 7; Mismatches 18; Indels 11; Gaps 1;

Qy 4 YGTNGVHCTKNSLWGVRLKNMKYDQNTTYMGRLODILLGWATGAFGKTFH 54
Db 331 YGT-----GRAVLFEVSYQSNWRYLEIVLYLIGRAGGYGALFN 370

RESULT 9
US-11-188-298-13006
; Sequence 13006, Application US/11188298
; Publication No. US2006007522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 13006
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(147)
; OTHER INFORMATION: unsure at all xaa locations
US-11-188-298-13006

Query Match      18.4%; Score 56; DB 11; Length 147;
Best Local Similarity 32.7%; Pred. No. 4.2;
Matches 17; Conservative 10; Mismatches 23; Indels 2; Gaps 2;

Qy 3 YYGTNGVHCTKNSLW-GKVRLLKNMKYDQNT-TYMGRLQDILLGWATGAFGKT 52
```

```
Db 26 YETTELVHENRNGTYQWIVKPKSVNYQFKTINTHPKLGVLVGVGNGNST 77

RESULT 10
US-11-188-298-15017
; Sequence 15017, Application US/11188298
; Publication No. US2006007522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 15017
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Glycine max
US-11-188-298-15017

Query Match      18.3%; Score 55.5; DB 11; Length 226;
Best Local Similarity 36.5%; Pred. No. 7.9;
Matches 19; Conservative 9; Mismatches 21; Indels 3; Gaps 3;

Qy 3 YYGTNGVHCTKNSLW-GKVRLLKNMKYDQNT-TYMGRLQDILLGWATGAFGKT 52
Db 26 YETTELVHENRNGTYQWIVKPKSVNYQFKTINTHPKLGVLVGVGNGNST 76

RESULT 11
US-10-915-002-299
; Sequence 299, Application US/10915002
; Publication No. US20060078950A1
; GENERAL INFORMATION:
; APPLICANT: Proguleske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTI
; TITLE OF INVENTION: USE IN DIAGNOSIS ANTIGENS FOR USE IN THE DIAGNOSIS, TREATMENT, A
; TITLE OF INVENTION: PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/10/915,002
; CURRENT FILING DATE: 2004-08-10
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 299
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-915-002-299

Query Match      17.9%; Score 54.5; DB 9; Length 480;
Best Local Similarity 28.2%; Pred. No. 25;
Matches 11; Conservative 12; Mismatches 13; Indels 3; Gaps 2;

Qy 3 YYGTNGVH--CTKNSL-WGKVRLLKNMKYDQNTTYMGRLO 38
Db 296 FYAKSGMHFYATKHNFEWDEKLSIKGMLFDRDKDLEGKMK 334

RESULT 12
US-10-921-415-3
; Sequence 3, Application US/10921415
; Publication No. US20060014681A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Xiaojiang
; APPLICANT: Holers, V. Michael
; TITLE OF INVENTION: THREE-DIMENSIONAL STRUCTURE OF COMPLEMENT RECEPTOR TYPE 2 AND US
; TITLE OF INVENTION: THERBOF
; FILE REFERENCE: 2848-43
```

